Autonomous Auscultation of the Human Heart

by

J.S.F. Botha

Thesis presented at the University of Stellenbosch
in partial fulfilment of the requirements for the
degree of

Master of Science in Mechatronic Engineering

Department of Mechanical and Mechatronic Engineering
University of Stellenbosch
Private Bag X1, 1702 Matieland, South Africa

Supervisor: Prof. C. Scheffer

March 2010

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature: .....

J.S.F. Botha

Date: 2009/12/07

ABSTRACT

The research presented in this thesis serves to provide a tool to autonomously screen for cardiovascular disease in the rural areas of Africa. Vital information thus obtained from patients can be communicated to advanced medical centres by Telemedicine. Cardiovascular disease is then detected in its initial stages, which is essential to its effective treatment. The system developed in this study uses recorded heart sounds and electrocardiogram signals to distinguish between normal and abnormal heart conditions. This system improves on standard diagnostic tools in that it does not require cumbersome and expensive imaging equipment or a highly trained operator.

Heart sound- and electrocardiogram signals from 62 volunteers were recorded with the prototype Precordialcardiogram device as part of a clinical study to aid in the development of the autonomous auscultation software and to screen patients for cardiovascular disease. These volunteers consisted of 28 patients of Tygerberg Hospital with cardiovascular disease and, for control purposes, 34 persons with normal heart conditions.

The autonomous auscultation system developed during this study, interprets data obtained with the Precordialcardiogram device to autonomously acquire a normal or abnormal diagnosis. The system employs wavelet soft thresholding to denoise the recorded signals, followed by the segmentation of heart sound by identifying peaks in the electrocardiogram. Novel frequency spectral information was extracted as features from the heart sounds, by means of ensemble empirical mode decomposition and auto regressive modelling. These features proved to be particularly significant and played a major role in the screening capability of the system. New time domain based features were identified, established on the specific characteristics of the various cardiovascular diseases encountered during the study. These features were extracted via the energy ratios between different parts of ventricular systole and diastole of each recorded cardiac cycle.

The respective features were classified to characterise typical heart diseases as well as healthy hearts with an ensemble artificial neural network. Herein the decisions of all the members were combined to obtain a final diagnosis. The performance of the autonomous auscultation system used in concert with the Precordialcardiogram device prototype, as determined through the leave-one-out cross-validation method, had a sensitivity rating of 82% and a specificity rating of 88%. These results demonstrate the potential benefit of the Precordialcardiogram device and the developed autonomous auscultation software in a Telemedicine environment.

SINOPSIS

Hierdie tesis beskryf die navorsing van 'n outonome toetsing en sifting stelsel vir kardiovaskulêre siektes in landelike dele van Afrika, vanwaar mediese inligting per telefoon versend kan word. Die apparaat maak vroeë opsporing van kardiovaskulêre siektes moontlik, wat essensieel is vir effektiewe behandeling daarvan en ook die koste-effek van hierdie siektes verminder. In die huidige ontwikkelde stelsel word normale sowel as abnormale hart-toestande getipeer met opnames van hartklanke sowel as elektrokardiogram-seine. Voordele wat hierdie stelsel bo standaard diagnostiese metodes het, sluit die hanteerbare formaat van die hele apparaat sowel as die nie-noodsaaklikheid van duur beeldskeppende apparaat, of hoogs opgeleide personeel.

Hartklank- en elektrokardiogramseine van 62 vrywilligers is met die prototipe "Precordialcardiogram" apparaat opgeneem om by te dra tot die ontwikkeling van die rekenaar sagteware vir die outonome auscultatsie stelsel en om die pasiëntsiftingsvermoë daarvan te toets. Die vrywilligers het 28 pasiënte van Tygerberg hospitaal met abnormale harttoestande ingesluit, sowel as ‘n kontrolegroep van 34 persone met normale harttoestande. Die outonome auskultasie-stelsel wat tot stand gekom het deur hierdie ondersoek maak gebruik van “wavelet” sagte drempeling om geraas uit die opgeneemde seine te verwyder. Daarna word die hartklanke gesegmenteer deur die pieke van die elektrokardiogram te identifiseer.

Deur middel van "ensemble empirical mode decomposition" en outoregressiewe modellering, is nuwe inligting aangaande die frekwensie spektra van hartklanke, aanwysend van spesifieke harttoestande, verkry. Die beduidendheid van hierdie eienskappe is bewys en het 'n belangrike rol in die siftingsvermoë van die stelsel gespeel. Hierbenewens is nuwe tyd-gebaseerde eienskappe van die onderskeie kardiovaskulêre siektes wat tydens die ondersoek bestudeer is, geïdentifiseer. Hierdie eienskappe is geëien deur die energie-verhoudings tussen verskillende dele van die ventrikulêre sistolie en diastolie van elke opgeneemde hartsiklus te ontleed.

'n "Ensemble artificial neural network" is gebruik om die geïdentifiseerde eienskappe van hartsiektes sowel as normale harttoestande, te klassifiseer. Hierin is besluite van al die lede van die netwerk gekombineer, ten einde ‘n finale diagnose te maak. Die klassifiseerder se geldigheid is kruis-bevestig deur middel van die laat-een-uit kruisbevestigings-metode.

Deur middel van die kruis-bevestigingsmetode is die bedryfsvermoëns van die outonome auskultasie-stelsel, toegerus met die "Precordialcardiogram" apparaat, repektiewelik op 82% vir sensitiwiteit en 88% vir spesifisiteit vasgestel. Hierdie

resultate demonstreer die benuttingspotensiaal van die apparaat in 'n Telemedisyne omgewing.

ACKNOWLEDGEMENTS

The author would like to thank Prof Cornie Scheffer for his unfailing leadership and support during the entire project and Dr Dirk Koekemoer from GeoAXon for providing the prototype precordialcardiogram device.

Thanks to Dr Wayne Lubbe for sourcing patients and performing the ECG examinations and Dr Helmut Weich for performing the auscultations. Thanks to Amanda, for her help with the final editing and proof reading of this thesis.

Lastly, the author would like to thank The Mathworks Inc. for its excellent engineering tools.

DEDICATIONS

To Adri

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

Units

dB	Decibels
Hz	Hertz
ms	Milliseconds

Abbreviations

ANN	Artificial neural network
AV	Atrioventricular
AR	Aortic regurgitation
AS	Aortic stenosis
CVD	Cardiovascular disease
ECG	Electrocardiograph
EEG	Electroencephalogram
EEMD	Ensemble empirical mode decomposition
EMD	Empirical mode decomposition
EMR	Electronic medical record
FFT	Fast Fourier Transform
FIR	Finite-duration impulse response
FT	Fourier Transform
HHT	Hilbert Haung transform
HMM	Hidden Markov models
IC	Inter costal
ICG	Impedance cardiograph
IMF	Intrinsic mode function
k-NN	k-nearest neighbour
MA	Moving average
MLP	Multi-layer perseptron
MR	Mitral regurgitation

MS	Mitral stenosis
MV	Mitral valve
PC	Personal computer
PCA	Principal component analysis
PCG	Precordialcardiogram
PDF	Probability density function
PR	Pulmonary regurgitation
PS	Pulmonary stenosis
RMS	Root mean square
ROC	Receiver operator characteristic
SA	Sinoatrial
SNR	Signal to noise ratio
SOF	Statistical overlap factor
STFT	Short time Fourier transform
SVM	Support vector machines
TR	Tricuspid regurgitation
USB	Universal serial bus
VSD	Ventricular septal defect

1. INTRODUCTION

Cardiac auscultation, the technique of listening to heart sounds with a stethoscope, has been the primary means of diagnosing cardiovascular disease (CVD) since the early 19th century. One of the first applications of modern auscultation has been in the aerospace industry where, since the 1980's, electronic stethoscopes were embedded in astronauts' suits to allow monitoring of their cardiac activity (Nicogossian et al., 2001).

The sounds produced by the human heart have been studied extensively and forms the basis of many clinicians' preliminary diagnosis. But, cardiac auscultation is a difficult skill to master and is limited by the clinician's ability to discern the components which constitute the heart sounds. In South Africa, let alone the rest of Africa, there are only a few specialists trained in the skill of auscultation. Auscultation also fails to provide a qualitative and quantitative measure with which to formulate a diagnosis (Visagie, 2007). Recently, auscultation is being complimented by phonocardiography with the aim of assisting physicians in their diagnosis. Phonocardiography is the recording of heart sounds with electronic stethoscopes and displaying their waveforms. However, heart sounds exhibit complex nonlinear and non-stationary characteristics which are not clearly discernible simply by viewing phonocardiogram images.

With the advent of advanced imaging technologies such as echocardiography, computed tomography and magnetic resonance imaging, clinicians may have further neglected the vital skill of auscultation. However sophisticated these imaging techniques may be, they are usually performed at a high cost and are not always available to patients, especially in rural areas or underserved communities.

The research presented in this thesis therefore aims to assist clinicians in the early detection of CVD. Through the development of a software system with the capability to automatically analyse and classify heart sounds either as normal or abnormal, this primary aim was accomplished. The developed system employed advanced signal processing and pattern recognition techniques, such as the empirical mode decomposition and ensemble neural networks, to autonomously screen patients for CVD. The system was tested in a clinical study where data was recorded with the prototype Precordialcardiogram (PCG) device, developed by GeoAxon Holdings, South Africa. Data was obtained from patients who suffered from CVD as well as from healthy participants, to serve as control for the experiment. All the participants were examined by a trained cardiologist and an echocardiograph, which is the commonly used gold standard for diagnosis of CVD, to verify their diagnosis and to validate the autonomous auscultation software system developed during this study.

Advantages related to the implementation of this system in a product such as

the prototype PCG device, will include major savings with regards to valuable specialist time through the avoidance of unnecessary referrals and the ability to automatically detect CVD in its early stages. This is paramount for its effective treatment to avoid serious long-term complications and permanent damage to the heart. Through the development of the autonomous auscultation software, this research will have far reaching consequences with regard to making modern healthcare more accessible to all patients.

1.1 Motivation and Background

Between 1997 and 2004, approximately 200 people died per day in South Africa because of some form of cardiovascular disease, with more than half of those occurring below the age of 65 years (Steyn, 2007). Premature deaths caused by CVD in people of working age (35-64 years) are expected to increase by 40% between 2000 and 2030 (Steyn, 2007). In 1991, the cost of CVD to the South African economy fell between R4.135 and R5.035 billion, which did not include the costs of rehabilitation and follow-up. This expenditure reflects 2%-3% of the gross domestic product or roughly 25% of all healthcare expenditure in South Africa (Pestana et al., 1996). According to Steyn (2007), early diagnosis and treatment of CVD is essential to lower the burden on the economy and improve patient survivability.

By comparison, the South African National survey for HIV, 2008, showed that 18% of all South African adults were living with HIV and that in 2006, 41% of all deaths in South Africa were as a result of AIDS. The cost of HIV and AIDS on the South African economy is estimated to be 17% of its gross domestic product by 2010 (Ministry of Health, 2009).

Many cardiovascular diseases cause murmurs and aberrations in heart sounds even before they produce other noticeable symptoms, such as physiological changes or changes in the electrocardiogram (Huiying et al., 1997). Detection of these sounds can thus aid in the early diagnosis of these types of cardiovascular diseases and result in more effective treatment by preventing permanent damage to the heart (Visagie, 2007).

Patients identified with cardiovascular disease are generally referred to a cardiologist. However, rural areas in South Africa currently suffer from the absence of such specialists. The medical personnel currently working in these areas, who are responsible for screening populations for cardiovascular disease, might not have the proper training or skills required for diagnosing such diseases through auscultation(Lubbe, 2009).

Telemedicine can be described as the use telecommunication technology to de-

liver healthcare services to sites that are physically distant from the host service provider. The Medical Research Counsel of South Africa has announced that the South African government is, “committed to providing basic healthcare to all South African citizens”, and to achieve this goal the government has, “identified Telemedicine as a strategic tool for facilitating the delivery of equitable healthcare services”.

The use of the PCG device as a screening tool to assist in the auscultation of patients will prove useful to healthcare workers in rural communities, by improving the early detection of cardiovascular abnormalities and preventing inappropriate referrals. The PCG device, with screening capabilities, has the potential to bring professional medical services closer to the rural population by sending recorded patient data via Telemedicine to a specialist for review (Lubbe, 2009).

1.2 Research Objectives

The objective of this research project is to develop software capable of detecting CVD from heart sounds recorded with electronic stethoscopes. In order to achieve this objective the following must be accomplished:

1. Record heart sounds and electrocardiogram (ECG) data from patients with CVD as well as from healthy participants.
2. Research and implement suitable signal processing algorithms for the detection of anomalies in the recorded heart sounds which would indicate the presence of CVD.
3. Develop an automated classification scheme that can successfully partition the healthy participants from those with CVD.
4. Cross-validate the implemented classification scheme in a clinical study to determine its capability to screen future patients.
5. Investigate the strengths, weaknesses, feasibility and the acceptance of the recording hardware by patients and clinicians in a hospital environment.

1.3 Thesis Outline

Chapter 2 starts with a background to the cardiac cycle, heart sounds and the electrical activity in the heart. Current literature and advances in the field of autonomous auscultation is also reviewed.

Chapter 3 discuss the clinical study conducted with the prototype Precordial-cardiogram (PCG) device. The previous generation hardware prototype, the Auscultation Jacket, and the data recorded with it are also briefly discussed. Thereafter, the data pre-processing steps are described.

In Chapter 4, the autonomous auscultation system implementation is discussed. The chapter includes the denoising and segmentation of heart sounds, as well as feature extraction and classification.

The results from cross-validation are presented in Chapter 5 and the conclusions and recommendations of the study are presented in Chapter 6.

2. LITERATURE REVIEW

This chapter gives an introduction to the human cardiovascular system, the functionality of the human heart and the sounds it produces. Background to the electrical system and measurement of the electrical activity within the human heart is also presented. A review of various signal processing techniques that has been used to analyse heart sounds is given and a comparison of different classification methods with respect to screening for CVD is made. The chapter concludes with a summary of the methods employed in this study and a comparison to previous studies is drawn.

2.1 The Cardiovascular System and the Cardiac Cycle

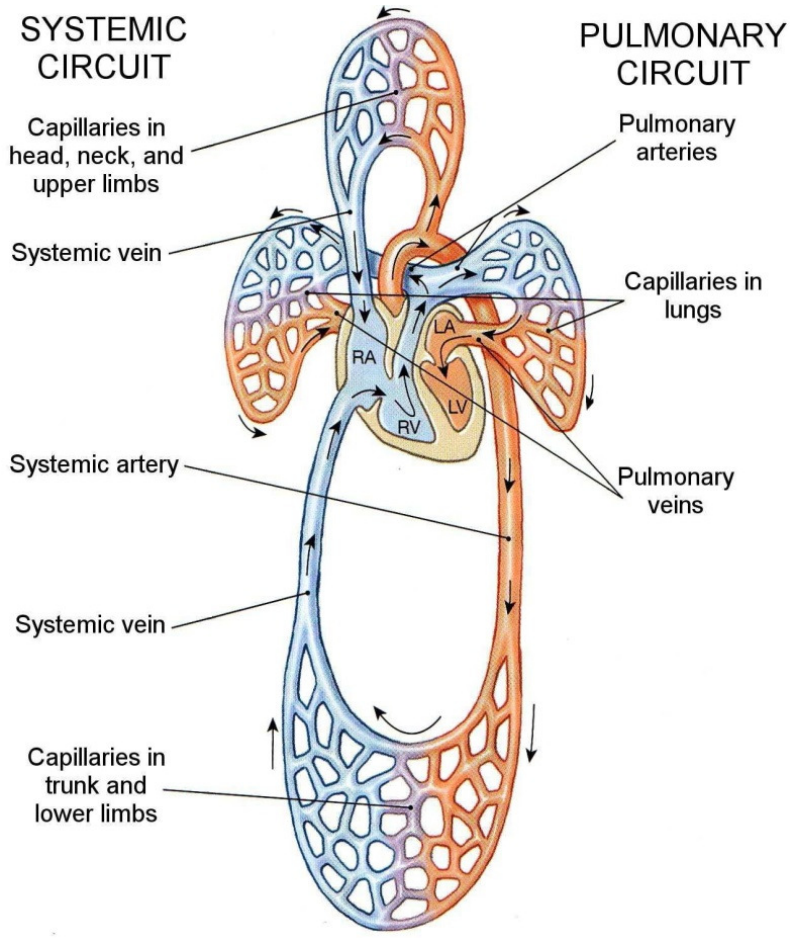

Figure 1: Cardiovascular circulatory system (Martini & Bartholomew, 2003).

The blood in the human body flows through a network of blood vessels (Figure 1) which connects the heart and peripheral tissue. The blood vessels can be categorised in two circuits. The pulmonary circuit carries deoxygenated blood from the heart to the lungs and return oxygenated blood from the lungs back to the heart. The systemic circuit carries oxygenated blood to the rest of the body and returns the deoxygenated blood back to the heart. The blood is propelled through these circuits in sequence, by the heart (Martini & Bartholomew, 2003).

The right atrium delivers deoxygenated blood to the heart from the systemic circuit (Figure 2). The right ventricle then pumps the blood to the pulmonary circuit during ventricular contraction. The left atrium collects oxygenated blood from the pulmonary circuit which it feeds to the left ventricle. The left ventricle then pumps the blood through the systemic circuit.

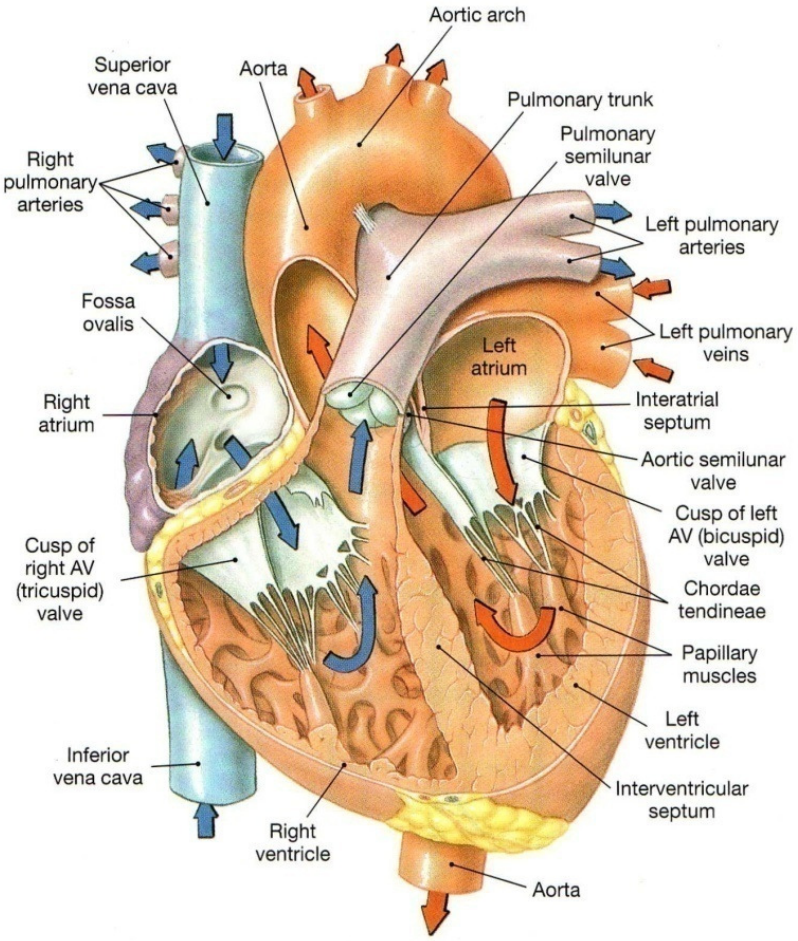

Figure 2: Frontal-section of the human heart (Martini & Bartholomew, 2003).

Four heart valves separate the atria and the systemic- and pulmonary circuits from the ventricles. The valves separating the atria and ventricles are the tricuspid and mitral (bicuspid) valves (Figure 2). The tricuspid valve consists of three cusps, while the mitral valve of only two. These cusps are connected to the inner cardiac muscle with tendons called the cordae tendineae. The purpose of the mitral and tricuspid valves is to prohibit the backflow of blood from the ventricles into the atria during ventricular contraction, but to allow blood flow when the ventricles are relaxed. Separating the ventricles and the pulmonary and systemic circuits are the pulmonary and aortic valves. These valves are also called semilunar valves, due to their shape. Their purpose is to prevent the backflow of blood into the ventricles from the pulmonary and aortic arteries when the ventricles relax after contraction (Martini & Bartholomew, 2003).

The cardiac cycle starts with atrial systole, when the atria contract to completely fill the ventricles with blood. Thereafter, atrial diastole and ventricular systole initiates simultaneously. During ventricular systole, the ventricles contract, forcing blood into the aortic and pulmonary arteries. When the ventricles relax after contraction, the ventricles and atria are mutually in diastole which lasts until atrial systole starts again, marking the start of a new cycle. The contractions of the atria are generally much weaker than that of the ventricles and not of as much concern. Therefore, in the remainder of this document, the terms systole and diastole are used to refer to ventricular systole and ventricular diastole, unless otherwise stated.

2.2 Heart Sounds and Auscultation

The sounds produced by the beating of the human heart are chiefly caused by the opening and closing of heart valves, as well as by the turbulence of the blood flowing through them. These sounds are named by a numbering convention which initiates at the start of systole. The two predominant sounds produced by a healthy heart can be described as “lubb-dubb” sounds and are called S1 and S2 respectively. Figure 4a) shows a heart sound recording (phonocardiogram) of a healthy volunteer, recorded with the PCG device, depicting S1 and S2. S1, heard at the start of systole, is largely caused by the closing of the mitral and tricuspid valves. S2 occurs at the onset of diastole and is mainly caused by the closure of the pulmonary and aortic valves (Martini & Bartholomew, 2003).

S3 is usually a low pitched sound and is prevalent in athletes with a slow heartbeat. In older patients, S3 usually indicates ventricular malfunction. S4 is caused by the contraction of the atria and the resulting movement of blood from the atria to the ventricles. A prominent S4 heart sound is usually caused by an enlarged or overgrown atrium pumping blood through a healthy tricuspid or mitral

valve (OSU Medical Centre, 2008). Extra sounds, indicating malfunctioning or diseased heart valves, include ejection sounds, clicks, snaps, rubs and rumbles.

Ejection sounds are usually caused by blood being pumped through a stenosed pulmonary or aortic valve. A stenosed valve is one which does not open properly and cause resistance to blood flow. In the case of mitral stenosis a diastolic rumble sound is usually heard after an opening snap. Clicks are extra sounds resulting from an elongation or rupture of the chordea tendons of either the tricuspid or mitral valves. The clicking sound is produced by the leaflets of such a valve moving into an atrium as the ventricles contract. Snaps are high-pitched sounds caused by the forward movement of a stenosed mitral or tricuspid valve as the ventricles relax during diastole. Pericardial rubbing sounds can be heard when the sack surrounding the heart becomes inflamed. The inflammation is usually the result of a viral infection and causes the wall of the heart muscle to rub against the pericardium.

Auscultation of the heart (the technique of listening to heart sounds), is usually performed at four specific positions on the thorax. These positions are situated where the sounds produced at each of the four heart valves are best heard and is also referred to by their respective heart valve names: tricuspid, mitral, pulmonic and aortic. Figure 3 shows the relative locations of the heart valves and their auscultation positions.

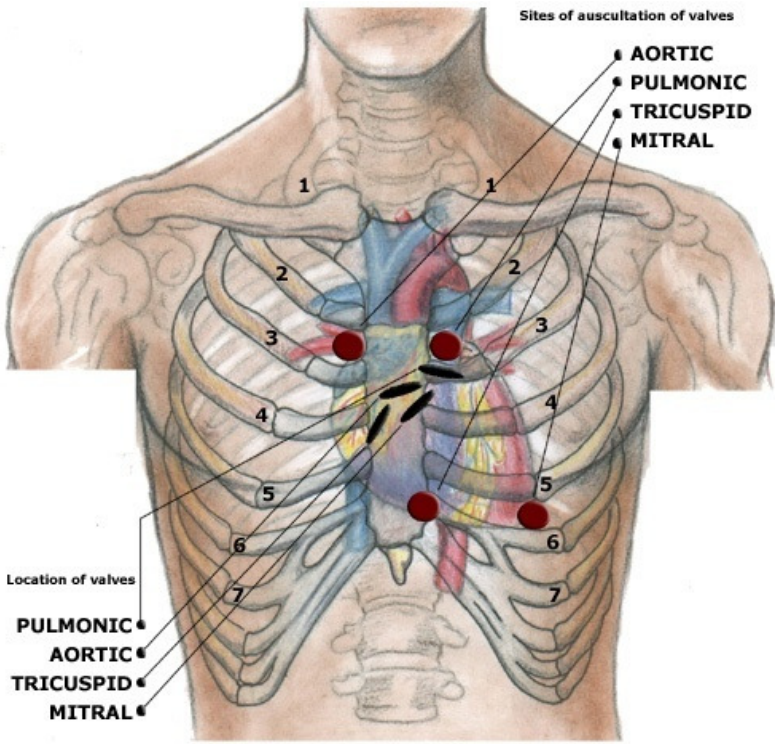

Figure 3: Heart valve and auscultation positions (Human, n.d.).

2.3 The Electrocardiogram

Contractions of the heart are controlled by a network of specialized muscle cells termed the conduction system of the heart. The conduction system, shown in Figure 5, initiates and distributes electrical impulses through the cardiac tissue. These electrical events, which cause the heart to contract, are detectable by measuring voltage potentials on the skin of a person's chest. A recorded time series depicting these voltages is called an electrocardiogram (ECG). Analysis of an ECG is useful in detecting cardiac arrhythmias and in the timing of cardiac cycles (Martini & Bartholomew, 2003).

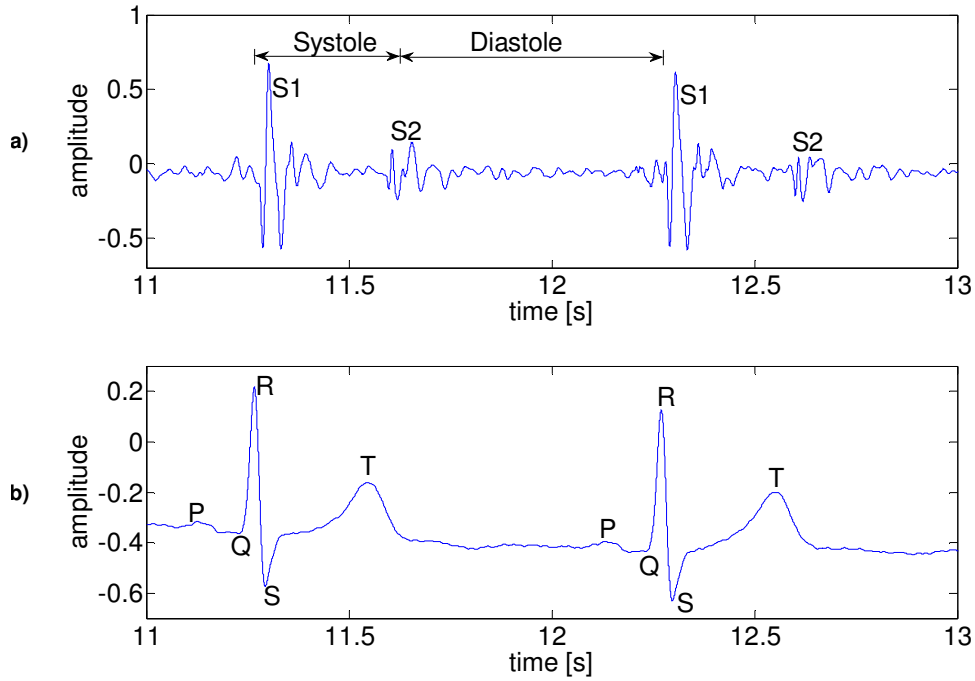

Figure 4: Synchronised recording of a healthy volunteer: a) Heart sound. b) ECG.

A typical ECG trace is shown in Figure 4b). The different waveforms numbered P to T represent the sequence of depolarization and repolarisation of the atria and ventricles. The P-wave represents the depolarisation of the atria. An action potential starts at the sinoatrial (SA) node (Figure 5) and propagates through the left and right atria causing them to contract simultaneously. The flat line between the P-wave and Q is the time the action potential travels through the atrioventricular (AV) node and the bundle of HIS.

The QRS-complex represents ventricular depolarisation. It contains the highest peak in the ECG trace since the ventricular muscle is much larger than that of the atria. Contraction of the ventricles starts shortly after the R-wave. The flat line segment between S and T is the time the ventricles remain depolarized and the T-wave indicates ventricular repolarisation. Atrial repolarisation is not apparent in the ECG trace since it occurs at the same time (and is masked by) the QRS-complex (Martini & Bartholomew, 2003).

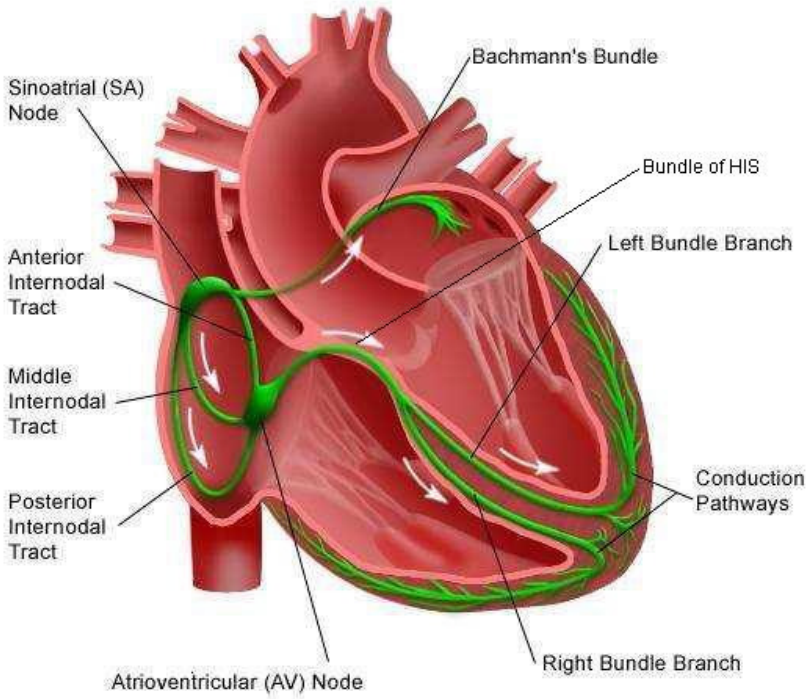

Figure 5: Electrical system of the heart (OSU Medical Centre, 2008).

2.4 Previous Research

A summary of previous research in the fields constituting autonomous auscultation is presented in this section. These include the denoising of heart sounds, segmentation of heart sounds, signal processing methods applicable to the study of heart sounds and classification schemes based on heart sound features. Finally, a comparison between previous research and this study is drawn.

2.4.1 Denoising

Signal contamination is a phenomenon present in most physical measurements. Its effect is to corrupt the data being measured by introducing false and unwanted information (noise) to the data. The presence of noise in any measurement is frequently related to two causes: 1) Environmental effects unrelated to the sampling procedure, and 2) constraints owed to the physical nature of the sampling hardware or the recording procedure. The effects of noise can often be reduced through software filters and denoising algorithms such as those discussed in the following paragraphs.

Adaptive filters have been widely used in applications where the statistical characteristics of the signals to be denoised are either unknown or time variant (non-stationary), such as in the case of heart sounds. Two classes of algorithms are frequently used to calculate the adaptive filter coefficients, 1) the least-means-square algorithm and 2) the recursive least-squares algorithm (Proakis & Manolakis, 2007). Multireference adaptive noise cancelling has been successfully used to improve the signal to noise ratio (SNR) of electroencephalograms (EEG) (James et al., 1997). Mittra et al. (2007) compared various denoising techniques for foetal phonocardiograms and concluded that adaptive filtering using the recursive least-squares algorithm is an appropriate methodology for these signals (Mittra et al., 2007).

Wavelet threshold denoising has been used extensively to denoise heart sound recordings. The heart sounds are decomposed into wavelet coefficients of which some are removed using a thresholding strategy. The remaining coefficients are recomposed to yield the denoised heart sound. The parameters for wavelet denoising are the chosen wavelet family, the level at which to perform the decomposition and the method used for calculating the threshold. The performance of different types of wavelets for the purpose of denoising heart sounds has been evaluated by Messer et al. (2001). Their research concluded that wavelets from the Debouche families decomposed at the 5^{th} level produced the best denoising results when applied to heart sounds.

2.4.2 Segmentation

The purpose of heart sound segmentation is to separate all the cardiac cycles in a recording so that each can be analysed individually. Three methods are currently employed to segment heart sounds. The first is to manually mark the start of each cycle by listening to the sounds. This is typically done by a cardiologist trained in cardiac auscultation.

The second involves applying sophisticated algorithms to the heart sounds. These algorithms are designed to automatically recognise recurring sounds in the recording which it uses as the basis for segmentation. The third method is to record an ECG or use a heartbeat monitor synchronous to the heart sound recording. The QRS peaks in the ECG are detected using a peak detection algorithm and is used to mark the start of each cardiac cycle.

The start of systole can be identified effectively using an ECG synchronised to the heart sound recording, but the start of diastole in the cardiac cycle then still needs to be determined. Malarvilli et al. (2003) computed the instantaneous frequencies of ECG recordings to segment heart sounds and identify S1 and S2. They found the method to be effective in segmenting their data which consisted of 210 recorded abnormal cardiac cycles (Malarvili et al., 2003). Rangayyan (2002) proposed a method for segmenting heart sounds by applying a first derivative operator and moving average (MA) filter to the ECG (Rangayyan, 2002).

Automatic segmentation of heart sounds without an ECG has previously been implemented in autonomous auscultation systems through the application of complex signal processing techniques. Torry and Mood (1995) implemented auto regressive modelling to estimate the power spectral densities and energies in certain frequency bands to segment heart sounds. In their study they applied their method to the segmentation of 30 patients' heart sounds and found that the segmentation agreed largely with that performed by a cardiologist, who listened to the heart sounds (Torry & Mood, 1995). Liang et al. (1997) based segmentation on the normalized Shannon energy of the phonocardiogram. Their method correctly segmented 93% of 37 patient recordings (Liang et al., 1997). Gupta et al. (2005) proposed a novel method for segmenting heart sounds by using homomorphic filtering and k-means clustering and reported good segmentation performance (Gupta et al., 2005). Gamero & Watrous (2003) segmented heart sounds by modelling systole and diastole using hidden Markov models as probabilistic finite state-machines. They evaluated their segmentation method with the phonocardiograms of 80 patients and obtained detection sensitivity for S1 and S2 of 93% and a detection positive predictive value of 97% (Gamero & Watrous, 2003).

2.4.3 Signal Processing

Many physical signals produced by natural phenomena such as the human heartbeat are best characterized statistically as random processes. Heart sounds exhibit extremely nonlinear and non-stationary characteristics and can only be observed for a finite duration (Visagie, 2007). The effect of this is that the length of the recording generally determines the time variation of the signal statistics. These characteristics, exhibited by heart sounds, make them a challenge to analyse in an

automated way and signal processing techniques that preserves time and frequency information needs to be employed (Reed et al., 2004). Some of the prominent methods, generally regarded to be suitable for analysing signals of this type, include the short time Fourier transform (STFT), wavelet analysis and, the recently proposed empirical mode decomposition (EMD) (Ari & Saha, 2008).

The STFT applies traditional Fourier spectral analysis to a finite, short time window of the data. The window is then slid across the data record to obtain a time-frequency distribution. This method is easily implemented and makes use of the well-defined and popular Fourier transform. However, a disadvantage of this method is that it assumes the data to be piece-wise stationary. This has been shown to rarely be the case with heart sounds revealing abnormal cardiac conditions (Reed et al., 2004). Sengur & Turkoglu (2008) extracted features based on the STFT from heart sounds and have shown them to be capable of accurately separating normal and certain abnormal heart conditions through classification with their artificial neural network (Sengur & Turkoglu, 2008).

Wavelet analysis is a non-parametric spectral analysis method which allows for an adjustable windowing length, as opposed to the STFT. It has proved to be very useful in analysing data which exhibit gradual frequency changes, but it suffers from leakage due to the limited length of the wavelet function. Debbal and Bereski (2004) were able to identify the aortic and pulmonary components in the second heart sound along with their frequency range and localization using wavelet analysis (Debbal & Bereksi, 2004). Ning et al. (2009) used wavelets and auto regressive modelling to analyse heart sounds and showed that it provided a quantitative means with which to delineate systolic murmurs (Ning et al., 2009).

EMD is a method that has been designed for analysing complex data such as those typically representing natural phenomena. Originally the method was used for studying ocean currents, but because of its advantages its application has spread to other fields, including the analysis of heart sounds. The most applicable advantage of the method is that it is able to uniquely decompose a non-linear, non-stationary time series into a sum of oscillatory components. The decomposition method is adaptive and based solely on the signal itself, contrary to methods employing Fourier transforms whose basis functions are linear combinations of fixed sinusoids (Huang et al., 1998). Charleston et al. (2006) has shown EMD to be capable of separating native heart sound components by, first validating the method on simulated heart sounds, and then testing it on real heart sound recordings (Charleston et al., 2006). Echeverria et al. (2001) has shown EMD to be able to provide useful features from heart sounds for the study of heart rate variability (Echeverria et al., 2001). The EMD method has also been implemented for the purpose of autonomous auscultation. Ari and Saha (2008) compared features extracted from heart sounds using EMD and wavelet analysis by classifying the features with a grow-and-learn network and an artificial neural network. They re-

ported that the EMD-based features generally resulted in better classification performance between normal and abnormal heart sounds, than the corresponding wavelet based features. In their study, they reported an overall classification accuracy of 95% with the EMD features and artificial neural networks (Ari & Saha, 2008).

2.4.4 Classification

Artificial neural networks (ANN) have been widely implemented for the purpose of classifying heart sounds. They are able to model complex nonlinear systems by using parametric forms for its basis functions, in which the parameters are adapted during training. The most successful type of ANN, in the context of pattern recognition, is the feed-forward neural network, also known as the multilayer perceptron (Bishop, 2006). Other notable methods that have successfully been implemented to classify heart sound features are support vector machines and hidden Markov models. A comparison of classification results and methods that has been reported in studies involved in the field of autonomous auscultation is presented in the following paragraphs.

Gupta et al. (2005) implemented Debouchies-2 wavelet detail coefficients, decomposed at the 2^{nd} level, as features in their multilayer perceptron (MLP) neural network. The wavelet coefficients amounted to 32 features which they used to classify between normal and abnormal heart sounds. They reported a correct classification performance of 90% based on their heart sound data, which consisted of 230 abnormal and 110 normal cardiac cycles (Gupta et al., 2005). Sinha et al. (2007) performed wavelet decomposition of phonocardiograms with a Coifman 4^{th} order wavelet kernel. They used the coefficients as features in an ANN classification system to detect signs of mitral regurgitation in the recordings and obtained a correct classification performance of 95% in their clinical study (Sinha et al., 2007). Turkoglu et al. (2002) extracted features from Doppler heart sounds using wavelet transforms and the STFT. They implemented the features in an ANN for classification and obtained correct classification rates of 94% for normal participants and 96% for abnormal patients (Turkoglu et al., 2002). Leung et al. (2000) used the trimmed-mean-spectrogram time-frequency analysis method to extract features from recorded phonocardiograms. They implemented the features in an ANN and tested the system in a clinical trial consisting of 18 normal participants and 37 abnormal patients. Their system was able to achieve a sensitivity of 97% and a specificity of 94% (Leung et al., 2000). Andrisevic et al. (2005) developed an autonomous auscultation system with an ANN classifier. They used principal component analysis (PCA) and block processing to prepare features for the network. They obtained a sensitivity of 64.7% and specificity of 70.5% by screening for CVD with their system (Andrisevic et al., 2005).

Ensemble artificial neural networks have been shown by previous research to give more accurate results under certain conditions compared to unitary networks. The objective of an ensemble of neural networks is to combine the individual outputs of each ensemble member to obtain better generalisation. Sharkey (1996) first published an overview on the research of combining artificial neural networks and the fundamental challenges involved with such an approach. Some of the challenges include combining the outputs of ensemble members and methods used for creating candidate ensemble members (Sharkey, 1996). Dietterich (2000) showed in a study of machine learning methods how ensembles of classifiers can have remarkable advantages over single classifiers (Dietterich, 2000). Zhou et al. (2001) used an ensemble artificial neural network to successfully identify lung cancer cells in images. Gilbert & Tan (2003) compared ensemble classification methods to single classifiers and found that the ensemble often produced better results than single classifiers in their study of detecting cancer using gene expression data. Liu et al. (2004) implemented ensemble artificial neural networks to also classify gene expressions and found that it considerably improved accuracy and robustness compared to their previous, single ANN. Bhatikar et al. (2004) developed a screening system for CVD in paediatrics based on a decision-of-experts ensemble ANN classification scheme. They have shown, through a clinical study, that their system obtained a sensitivity and specificity of 88% and 83% respectively.

Other schemes employed in the classification of heart sounds include hidden Markov models (HMM), k-nearest neighbours, support vector machines (SVM) and linear discriminant models (Visagie, 2007). Sengur & Turkoglu (2008) developed a fuzzy k-NN heart sound classification system to detect heart valve disease and obtained a sensitivity and specificity of 95.9% and 96% respectively from testing 215 sample recordings. Chauhan et al. (2008) introduced a Mel-frequency ceptral coefficient to extract features from heart sounds. They used HMM to classify these features and obtained 90% classification accuracy for detecting diastolic murmurs.

2.5 Comparison of the Present Study with Previous Research

The study presented in this thesis follows directly from previous work done by Visagie (2007). However, the aim of this study was to use the new prototype Precordialcardiogram (PCG) device (Figure 6) to record data in a clinical environment and to use it as a platform for the development of the auscultation software to autonomously screen patients for CVD.

Synchronised stethoscope- and ECG signals from multiple sensors in the PCG device was employed in this study, similar to that of Visagie (2007), to record patient data. In contrast to many other methods reviewed, the PCG device procides the ability to observe a unique heart sound event at various auscultation positions, along with its characteristic ECG traces. Operation of the PCG device, unlike the auscultation jacket (Figure 9) used by Visagie (2007), is designed to be straightforward so that a trained nurse could quickly and effortlessly screen patients for CVD (Koekemoer & Scheffer, 2008). The PCG device and associated hardware is also designed to be low-cost to enable its swift deployment at many medical centres without the need for large capital investment. Table 1 gives a cost comparison between a standard ECG and echocardiograph (currently employed at medical centres) and an estimated cost of a commercial PCG device with embedded screening capabilities.

To develop and validate the performance of the autonomous auscultation system of this study, 60 participants' heart sounds and ECG signals were recorded at Tygerberg Hospital with the PCG device. Subsequently, the physical abilities of the device and its feasibility in a hospital environment were assessed and the performance of the entire system was determined. The performance of the system was calculated in terms of the sensitivity, specificity, accuracy, predictive values and likelihood ratios so that it can be accurately compared to previous studies. These measures are also used to describe the usefulness of the autonomous auscultation system as a diagnostic and screening tool. Additionally, the system is designed to give a novel confidence percentage on every diagnosis to indicate how conclusive a particular screening result is.

The software developed in this study denoises all the recorded heart sounds with wavelet soft threshold denoising. This method is similar to that used by Visagie (2007) as well as other studies on the denoising of heart sounds, as discussed in Section 2.4.1. In this study, the method was found to accurately denoise the heart sounds without removing significant oscillations.

Segmentation of the heart sounds was performed with the assistance of the internal ECG in the PCG device. The segmentation algorithm takes into account the time history of successive identified peaks to assist in the elimination of falsely identified QRS-peaks in the ECG. The method proved extremely robust and performed successful segmentations even on very noisy ECG traces, such as illustrated in Figure 17 and Figure 22.

In the present study, novel use of the ensemble empirical mode decomposition was made to study heart sounds and analyse them in an automated way. With its ability to identify significant oscillation modes, the developed software can uniquely use auto regressive modelling to extract features for use as input to the ensemble ANN classification scheme. Ensemble neural network decisions are then combined to obtain a diagnosis as well as a confidence percentage on that

diagnosis. This ability gives the autonomous auscultation system developed in this study a clear advantage and makes it a much more valuable tool for providing decision support to clinicians when compared to many of the previously proposed systems.

Table 1: Cost comparison of cardiac diagnostic resources (Koekemoer, 2008).

	Echocardiograph	**ECG**	**PCG device**
Capital Expense	R500,000	R20,000	R40,000
Operator	Sonographer	Nurse	Nurse
Salary expense	R160,000	R130,000	R130,000

2.6 Chapter Summary

This chapter presented an introduction to the cardiovascular system and the features of heart sounds. It gave a review on various signal processing methods previously used to study and classify heart sounds for the purpose of screening for CVD. Finally, the main differences and similarities between the present study and previous studies were discussed.

3. HARDWARE AND DATA AQUISITION

This chapter describes the Auscultation Jacket and Precordialcardiogram (PCG) prototype devices that have been used to record heart sounds and ECG data. The clinical study performed at Tygerberg Hospital to record data with the PCG device is described and preparation of the recorded data, before it is submitted to the auscultation software, is explained. The chapter concludes with the development of a single-lead ECG, which samples data through a universal serial bus (USB) mini-soundcard, to extend the capability of the PCG device.

3.1 Precordialcardiogram device

The PCG prototype is a handheld device produced for this study by GeoAxon Holdings Pty. Ltd. It is designed to fit the average adult's chest to record ECG and heart sound signals (Koekemoer & Scheffer, 2008). It has a commercial, 12-lead ECG (IQTec, Cape Town, South Africa) and four electronic stethoscopes embedded inside the plastic housing. The device connects to (and is powered by) a personal computer (PC) via a USB hub and two USB extension cables. The PC controls the recording of signals and stores the recordings on its hard drive for analysis by the autonomous auscultation software.

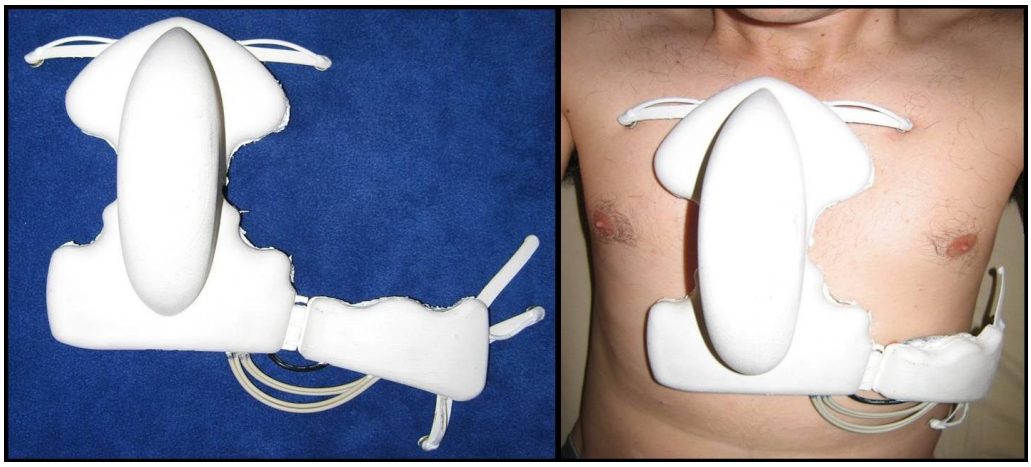

Figure 6: Prototype Precordialcardiogram device, top view (photos: JSF Botha).

Figure 6 shows the PCG device placed over the chest such that a recording can be made. The lateral part of the device can hinge separately to the main embodi-

ment, allowing the device to snugly fit most adults' chest circumferences (Koekemoer & Scheffer, 2008). The ECG electrodes can be seen in Figure 7 whilst the stethoscopes and diaphragms, housed beneath green silicone jelly, cannot be seen directly. Figure 8 shows the inside of the PCG housing. The pulmonic and aortic electronic stethoscopes can be seen on the left of the figure. The IQtec ECG and Soundcard-ECG is mounted vertically in the centre of the device and the USB soundcard and USB hub is on the right.

The stethoscopes sample at a resolution of 16-bits and a frequency of 8 kHz, whilst the ECG samples electrical potentials on the skin at 500 Hz. The ECG and stethoscope signals are synchronised at the start of each recording by tapping the device to create intentional artefacts. These artefacts then serve as reference to synchronise the signals manually after recordings were made. The location of the stethoscopes in the PCG housing is designed so that it coincides with the regular auscultation positions shown in Figure 3 (Koekemoer & Scheffer, 2008).

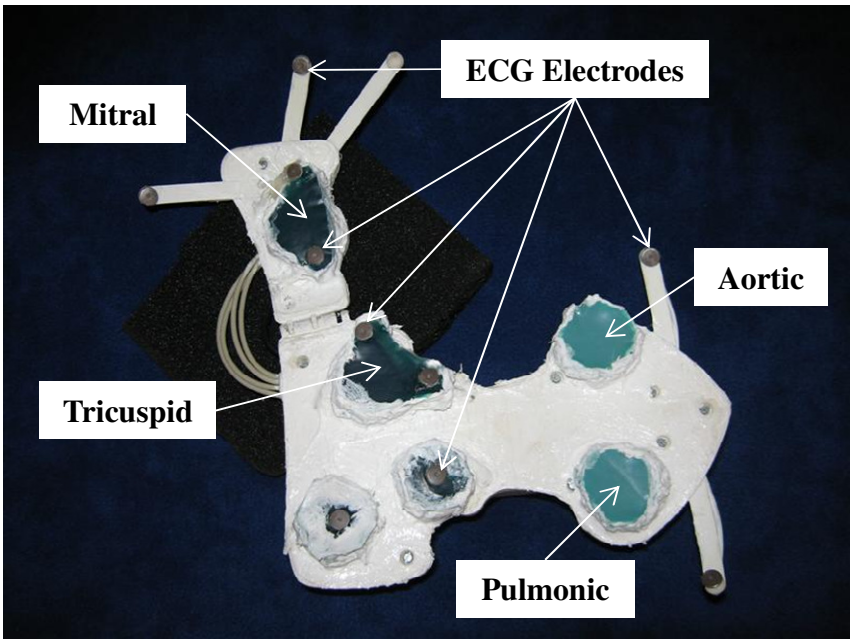

Figure 7: Prototype Precordialcardiogram device showing stethoscope positions and ECG electrodes (photo: JSF Botha).

Recordings of 60 volunteers were made at Tygerberg Hospital as part of the clinical study of this research project. The volunteers comprised of 30 healthy participants for experimental control and 30 patients with some or other cardiovascular disease that is detectable through auscultation. Unfortunately, two of the patients with abnormal heart conditions and five of the normal participants' re-

cordings could not be synchronised because of poor ECG quality. Subsequently, their data could not be used to evaluate the classification system. The causes underlying the poor ECG recording quality are discussed further in Section 6.2. To make up for the loss of data, nine more volunteers with normal heart conditions were recorded. The total amount of usable data collected with the PCG device therefore amounted to 28 abnormal (with heart disease) and 34 normal (no heart disease) participants with 3 to 4 minutes of ECG and heart sounds each.

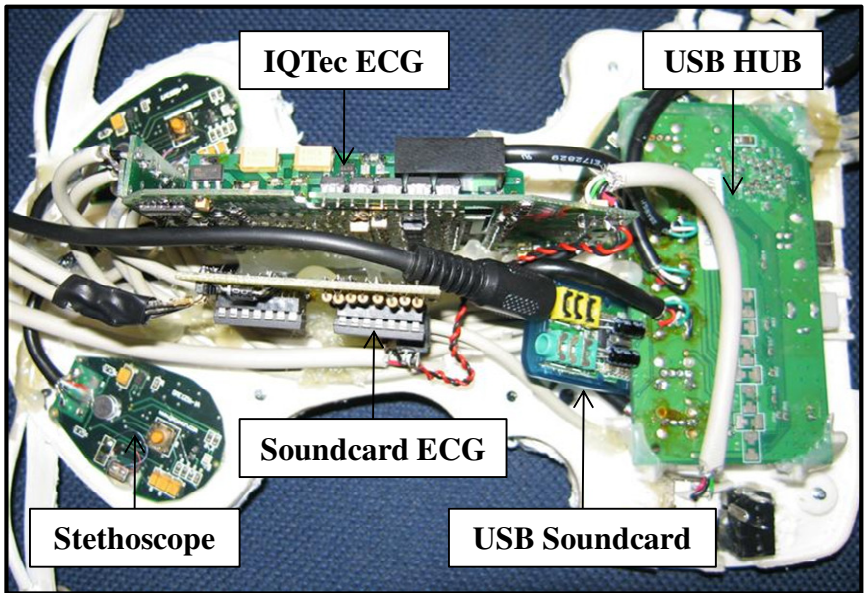

Figure 8: PCG device showing electronic stethoscopes, IQTec ECG, soundcard ECG, USB soundcard and USB hub (photo: JSF Botha).

3.2 Auscultation Jacket

The Auscultation Jacket used in the research of Visagie (2007), is designed to record heart sounds, lung sounds, ECG and impedance cardiogram (ICG) signals simultaneously from several sensors. The sensing elements consist of electronic stethoscopes for the sound recordings and electrodes for the ECG and ICG.

The body of the Auscultation Jacket is made from flexible neoprene foam and is designed to fit the patient such that the position of the electronic stethoscopes coincides with the normal auscultation locations (Koekemoer & Scheffer, 2008). Figure 9 shows the Auscultation Jacket fitted to a mannequin to illustrate how it would fit a patient. The fitment of the neckpiece housing the ICG electrodes is also shown.

Figure 9: Auscultation Jacket fitted over mannequin (Koekemoer & Scheffer, 2008)

The location of the stethoscopes inside the neoprene is shown in Figure 10. The stethoscopes sample data at a resolution of 16-bits and a frequency of 2 kHz whilst the ECG and ICG sample at 500 Hz (Koekemoer & Scheffer, 2008).

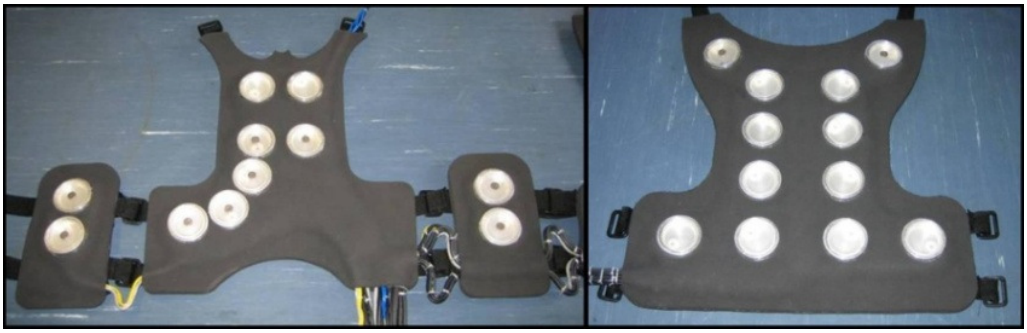

Figure 10: Auscultation Jacket showing stethoscopes and ECG electrodes (Koekemoer & Scheffer, 2008)

Participant recordings performed with the Auscultation Jacket in 2005 (Visagie, 2007) were used additional to the data collected during the present study. The recordings, made with the Auscultation Jacket, were of 21 patients who presented with cardiovascular disease and 31 healthy volunteers for experimental control. Due to the shape of the Auscultation Jacket, these recordings were performed exclusively on male volunteers (Visagie, 2007).

3.3 Clinical study at Tygerberg Hospital using the PCG device

For the clinical study, data from 62 participants was successfully recorded using the PCG prototype device at the TREAD research facility at Tygerberg Hospital. The procedure for performing recordings, outlined in the participant consent form in Appendix B, and summarised in Appendix C, was approved by the Committee for Human Research of the University of Stellenbosch (Nr. N07/02/026).

The purpose of using the PCG device to performing the recordings on patients was on the one hand to increase the size of the local heart sound and ECG database which, to date, only consisted of recordings made with the Auscultation Jacket. The recordings were also used to determine how accurately the implemented classification scheme was able to detect the presence of CVD in cross-validation tests and to evaluate the feasibility of such a system in a hospital environment.

The participants that were successfully recorded comprised of 37 male and 25 female volunteers and ranged from 13 years to 79 years of age. Twenty eight of the participants presented with CVD detectable through auscultation and 34 were normal, healthy volunteers for experimental control. Most of the abnormal patients presented severe cases of mitral regurgitation (MR), mitral stenosis (MS), aortic regurgitation (AR), aortic stenosis (AS), or ventricular septal defect (VSD). These abnormal conditions are discussed in more detail in Section 4.4.1 as part of the pathology-based analysis. Tables 2 and 3 show a summary of all the participants' information and the number of occurrences of each of the abnormal conditions. Most the abnormal patients presented combinations of these conditions with only a few conditions appearing on their own.

To verify the diagnosis of each participant's heart condition, each underwent an examination by a trained cardiologist using an echocardiograph. The echocardiograph is regarded as the gold standard in cardiac diagnostics and the results were used during the present study as a benchmark with which to compare the results of the developed autonomous auscultation system. A standard 12-lead ECG was used by Dr Lubbe (Lubbe, 2009), to record an electrocardiogram for each participant. All the participants had to sign the written consent form presented in Appendix B and the procedure for performing the recordings as listed in Appendix C, was followed for everyone. The details describing all the participants' information, the results of the echocardiograph and the conclusions of the cardiologist's examinations are listed in Table A4.

Three recording runs, also referred to as instances, were performed for each participant. The duration of each instance is approximately one minute or 70~90

cardiac cycles. During each instance the participant was instructed to inhale deeply and then to exhale, but not to inhale again for as long as possible. In a few cases the recording duration was shorter since cardiovascular disease caused a severe lack of stamina in some of the patients. To ensure that the length of the recordings was not taken into account by the classification system, the extracted features were averaged over all the recorded cycles of each participant.

Table 2: Summary of participants in clinical study.

	Maximum	**Minimum**	**Average**
Age	79 years	13 years	39 years
Weight	128 kg	45 kg	73 kg
Height	194 cm	152 cm	168 cm

Table 3: Occurrences of abnormal conditions in participants.

Abnormal Condition	**Number of occurrences**
Mitral regurgitation	11
Mitral stenosis	6
Aortic regurgitation	8
Aortic stenosis	10
Ventricular septal defect	2
Pulmonary stenosis	1

3.4 Data Preparation

Data preparation is necessary to save all the recordings into a standard format to enable the automated auscultation system to analyse and classify the data. This step has been deemed impractical to form part of the automated process since an open electronic medical record (EMR) standard does not yet exist for recording and saving synchronised ECG and heart sound signals.

The first task was to resample the ECG to the same sampling frequency as the

heart sound, 8 kHz in this case. One of the ECG channels, which recorded clearly, was chosen for segmentation and converted and saved in the same file format as the heart sounds. Intentional artefacts in the signals, created at the start of each recording, were then used to synchronize the ECG and heart sounds manually. Finally, the unusable portions in the recordings are removed and the result is saved in a 5-channel pulse-code modulated waveform audio format file. The first four channels in the file are the signals from the four electronic stethoscopes at the mitral, tricuspid, pulmonary and aortic auscultation positions (Figure 3) and the fifth is the ECG signal.

The recordings of each participant was organised by a naming convention. For example, the file name *DATA03_1_2.wav* refered to *participant identification number* 3, *instance* 1, *part* 2. The recording instance is a single recording run during which the participant was instructed to exhale and hold his or her breath. Some of the instances are broken down into parts (shorter recording strips) because they contained segments with excessive noise. The noise was primarily caused by people talking at a nearby examination or instruction is being given to the participant. These pieces of unusable recording are removed and the instance was consequently split into several parts.

3.5 Soundcard-ECG

To overcome the difficulties involved in synchronizing the commercial, 12-lead, IQTec ECG and stethoscopes built into the PCG device, an analogue ECG pre-amplifier circuit board, referred to as the Soundcard-ECG (Figure 11), was designed and built for the PCG device for use in future trials. The circuit diagram and printed circuit board layout of the Soundcard-ECG is presented in Appendix E. The Soundcard-ECG uses the analogue to digital converter of a small USB soundcard's microphone input to capture a single-lead ECG trace.

A raw recording of the Soundcard-ECG is shown in Figure 12b). The ECG is automatically synchronised with the heart sound, Figure 12a), by the recording software. The recording illustrates that it is feasible and practical to record a good quality ECG for heart sound segmentation purposes by using a low-cost USB sound card.

The USB soundcard and ECG is compact enough to fit inside the PCG device housing alongside the IQTec ECG, as shown in Figure 8. The major advantage of this device is that the stethoscopes and Soundcard-ECG are automatically synchronized by the audio recording software since the Soundcard-ECG now use the same *winsound*-device drivers as the electronic stethoscopes. The device is also cheap (less than R300), compact, and easy to assemble. The device is powered exclusively by the computer's USB port through which the data is also sampled.

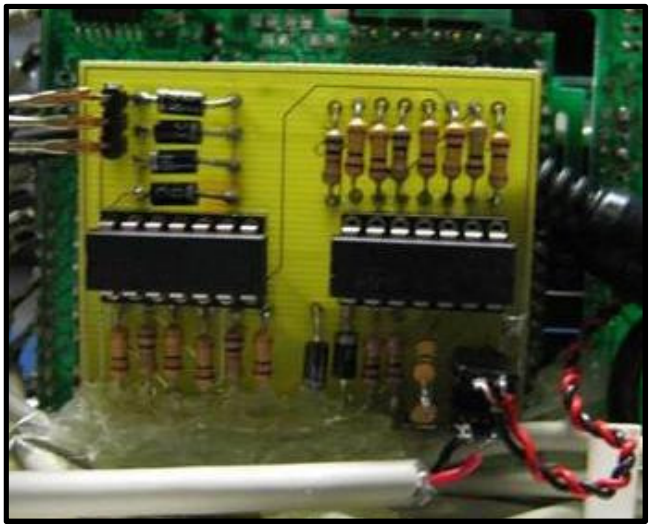

Figure 11: Soundcard-ECG (photo: JSF Botha).

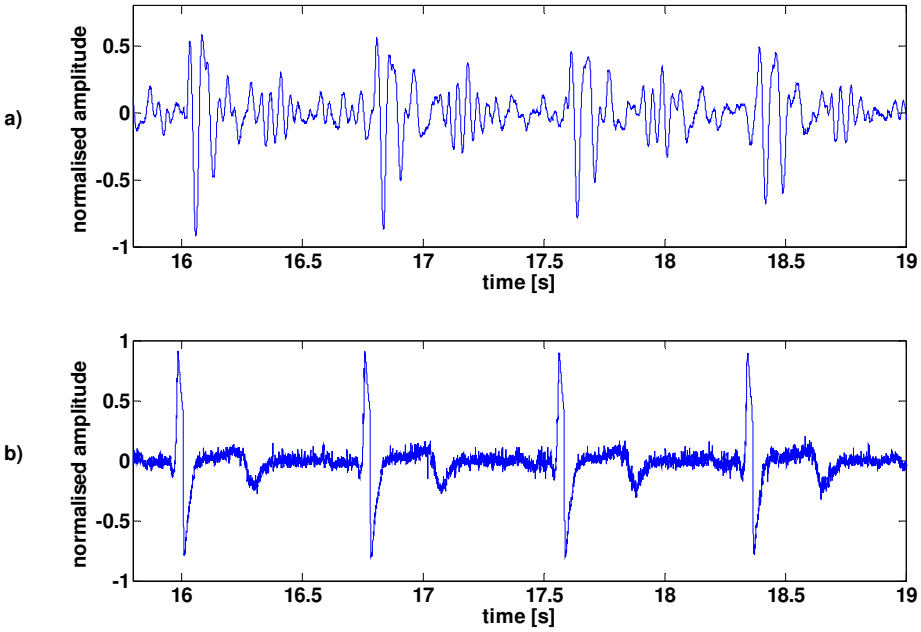

Figure 12: a) PCG heart sound recording. b) Raw Soundcard-ECG recording, automatically synchronised with sound in Figure 12 a).

3.6 Chapter Summary

The clinical study conducted with the PCG device at Tygerberg Hospital to screen patients for CVD is discussed. The components and functionality of the PCG device is outlined and the preparation of data that has been recorded with it, is described. In addition, the design and manufacture of an analogue ECG (Soundcard-ECG) which automatically synchronize with the heart sound recording, is explained.

4. METHODOLOGY

This chapter discusses the software implementation of the autonomous auscultation system. The system components chronologically send data to each other, as shown in Figure 13, to ultimately provide a normal (healthy heart condition) or abnormal (CVD present) diagnosis. The implementation was done in the Matlab (Mathworks Inc., Massachusetts, USA) environment with the addition of the signal processing-, wavelet-, neural network- and data acquisitioning toolboxes.

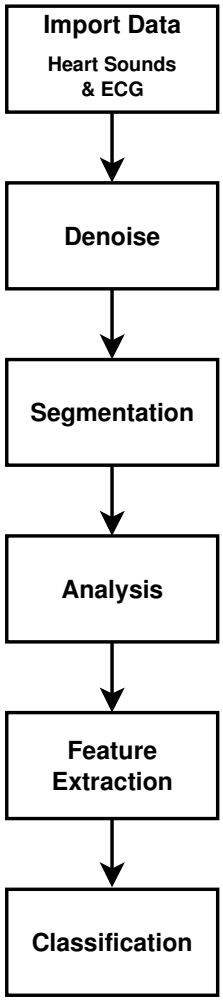

Figure 13: Overview of the autonomous auscultation system implementation.

The first component of the system imports the recorded heart sound and ECG signals of each participant into the Matlab environment and saves it in a structure whose fields are accessible to the other components. No other information regarding the participants, except their *participant identification number*, is given to the system. Thereafter, the data is denoised using wavelet soft thresholding followed by segmentation which splits the recordings into the systole and diastole cycles. Analysis of the heart sounds is then performed in the time- and frequency domain, followed by the extraction of features. The features are presented as input to the classifier which computes the probability of the heart sounds being either normal or abnormal.

4.1 Importing Participant Data

Data is presented to the autonomous auscultation software in the form of pulse code modulated waveform files where, for example, the file named, DATA03_1_2.wav, refers to participant number 3, instance 1, part 2. The file would contain five synchronised channels of which the first four is the four stethoscope recordings and the fifth is the ECG. All the data files of each participant is read and saved into a unique structure with the *wavread* function in Matlab.

As a final step, the amplitudes of all the stethoscope and ECG signals are individually normalised to the range $[-1 \quad 1]$. Therefore, all the figures representing signals in this thesis have a normalised amplitude projected on their y-axis.

4.2 Denoising the Heart Sound and ECG Signals

Signal contamination originating from environmental noise and the physical constraints of the recording equipment, made the heart sound and ECG signals unsuitable for direct analysis. This section describes the different types and sources of noise encountered while using the PCG device in the clinical study at Tygerberg Hospital. The methods used to prevent signal contamination and the implementation of software-based, automatic denoising is also discussed.

Noise resulting from the design of the PCG device is partly due to the fact that the electrodes and stethoscopes are fixed in position relative to a rigid housing. This produced a continuous shifting of the position of the electrodes and stethoscopes against the participants' chest, as it expanded and contracted during breathing. Because of this shifting, electrical interference was introduced to the ECG and the stethoscopes recorded "rubbing" noises.

To overcome these signal contaminations, ultrasound gel was placed on the participant's skin at the stethoscope locations before starting a recording. The gel reduced the friction between the skin and stethoscopes, but unfortunately only marginally reduced the rubbing noise. To address the electrical noise caused by the shifting of position of the ECG electrodes, an electrolyte solution was sprayed on the electrodes before placing the PCG device on the participant's chest. The electrolyte significantly increased skin-electrode conduction and reduced much of the electrical noise.

Environmental noise contributed significantly to the total magnitude of the noise present in the recordings. The major sources of environmental noise was people talking in the hallways, other examinations taking place close by, noise and electrical interference produced by other medical equipment in the examination room and general hospital background noise produced largely by air conditioners.

Active noise cancellation is incorporated in the electronic stethoscopes to reduce most of the effect of environmental noise in the sound recordings (Koekemoer, 2008). A second microphone on the opposite side of each stethoscope records environmental sounds which a microprocessor then inverts in its memory. The stethoscope signal and the inverted environment recording are combined, attenuating environmental noise. The active noise cancellation incorporated in the stethoscopes was successful in eliminating background speech from the recordings, but in some cases significant sounds such as the clicking sound of an artificial heart valve, was also attenuated.

Software-based denoising was further necessary for removing any remaining noise present in the recordings and to ensure that useful information can be extracted from them. The procedure followed to perform the software-based denoising is outlined in Figure 14.

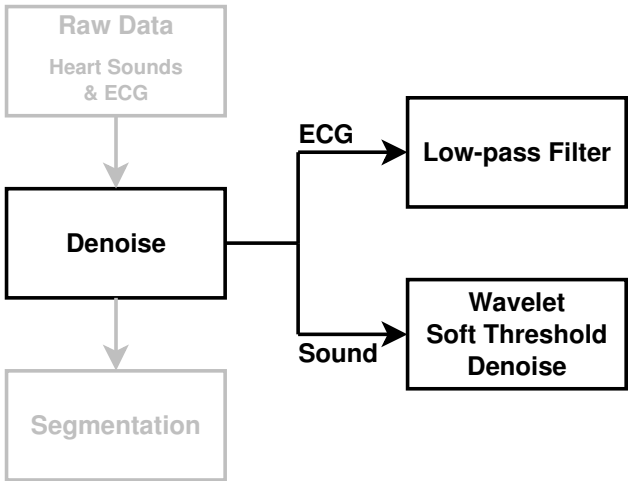

Figure 14: Denoising flow diagram.

Electrical interference in the ECG was found to be effectively removed through low-pass filtering. The interference is normally of a much higher frequency than the waveforms of interest (QRS-complex, T-wave etc.). A 6th order Butterworth low-pass filter with a -3dB cut-off frequency at 35 Hz was implemented and proved adequate for removing most of the electrical interference. The magnitude and frequency responses of the filter are shown in Figure 15 and its effective application to denoise a severely corrupted ECG (Figure 16) is illustrated by the result shown in Figure 17. The filter sufficiently denoised the ECG to allow the segmentation algorithm described in Section 4.3 to successfully divide the signal into systole and diastole cycles, as is also illustrated in Figure 17.

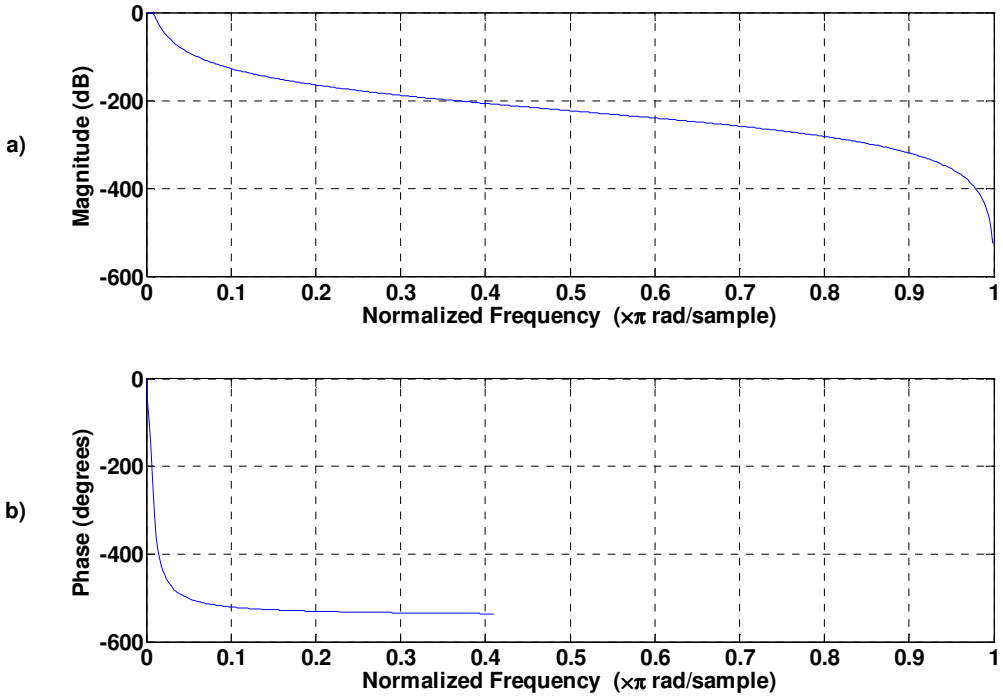

Figure 15: Phase- and magnitude response of the 6th order, low-pass Butterworth filter implemented to denoise the ECG signals.

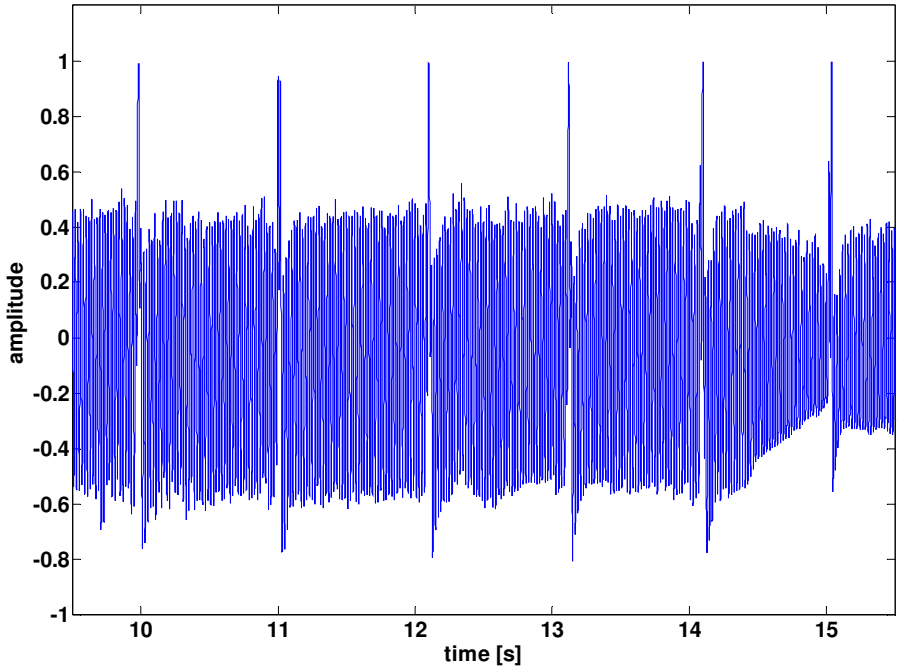

Figure 16: Unfiltered ECG recording.

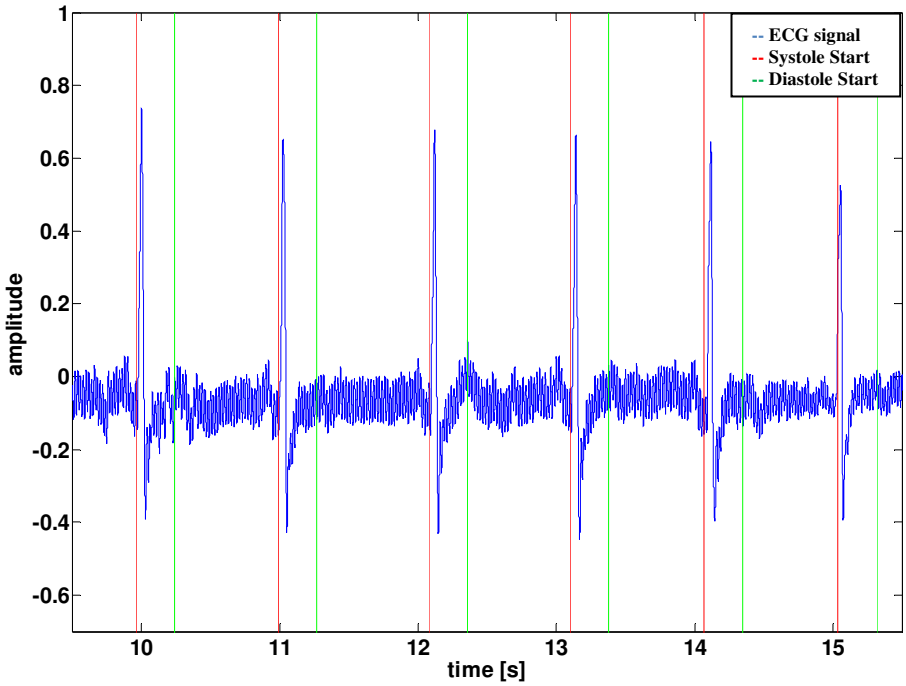

Figure 17: Filtered ECG recording of the example in Figure 16.

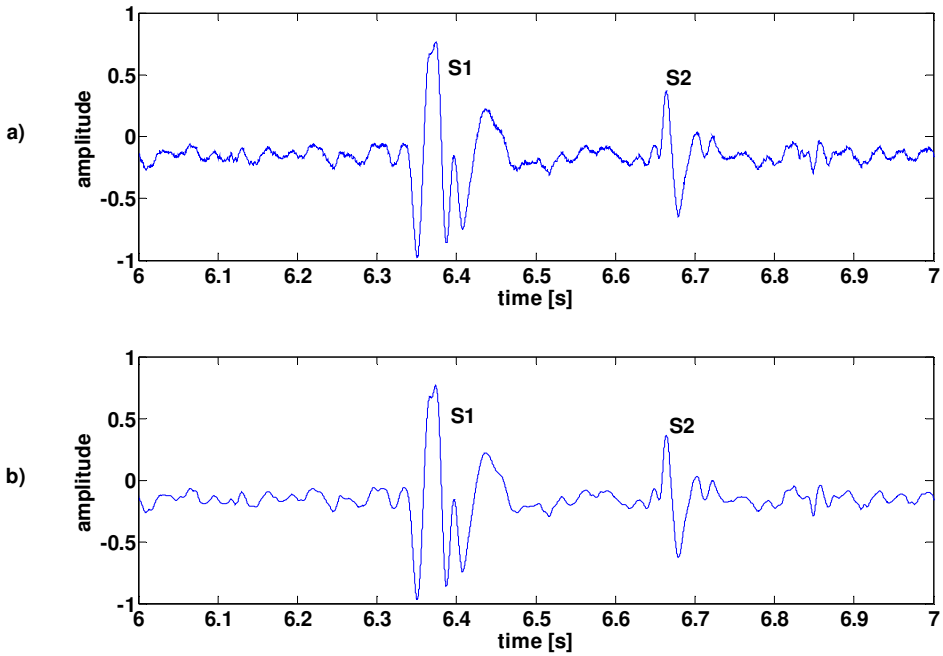

Figure 18: a) Recorded heart sound cycle. b) Denoised heart sound cycle.

Heart sound denoising was accomplished by decomposing the sound with a Debouche 7 wavelet at the 5^{th} level. Adaptive threshold selection was based on Steins unbiased risk estimate which minimizes the risk of threshold values given by a quadratic loss function. The threshold value was then further reduced by a factor of 0.3 to ensure that meaningful information was not removed from the signal. Soft thresholding was applied to all except the approximation wavelet coefficients. Finally, the coefficients were re-composed, yielding the denoised sound shown in Figure 18b).

4.3 Segmentation

Segmentation of heart sounds is the process by which a recorded sound is identified as either being part of systole or part of diastole of the cardiovascular cycle. Knowing at which time in the cardiovascular cycle a sound is produced is vital for the identification of components in that sound that may indicate a pathology. After the sound is segmented it is possible to isolate specific sound events during a specific heartbeat and compare them as they reappear in other cycles.

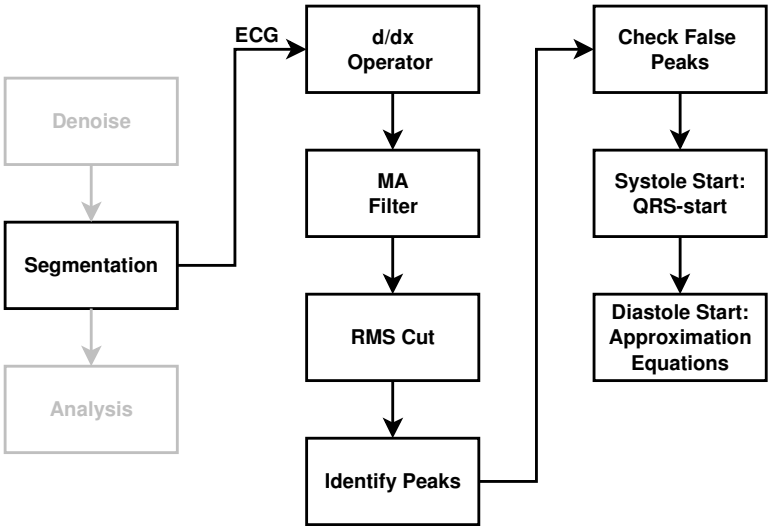

Figure 19: Segmentation flow diagram.

To facilitate segmentation an ECG was recorded and synchronised with the heart sounds. The ECG is used as a heartbeat monitor wherein the time between successive QRS-complexes are taken as entire cycles. The duration of each cycle is then used to estimate the start of diastole of that cycle through second order approximation equations.

The algorithm implemented to detect the QRS-complexes, outlined in Figure 19, is based on the methods described by Rangayyan (2002) and Visagie (2007) and uses a weighted, squared first derivative operator and a moving average filter. The first derivative operator acts as an edge detector as it accentuates areas in the signal which are steep and attenuates areas that are level. The MA filter only serves to smooth the resulting waveform.

First, the ECG is downsampled to 100 Hz to increase computational speed. If $x(n)$ is the ECG signal then,

$$y(n) = \sum_{i=1}^{N} | x(n-i+1) - x(n-i) |^2 \ (N-i+1) \tag{4.1}$$

represents the signal after applying the first derivative operator. N is the windowing length and the term, $(N-i+1)$, is the weighting factor. In this case, $N=1$ is used, similarly as in Visagie (2007). Rangayyan (2002) however, used a window width of $N=8$, but it was found that this resulted in excessive attenuation and incorrect detection of the start of systole.

After application of the first derivative operator, a MA filter is applied as a smoothing function to $y(n)$:

$$z(n) = \frac{1}{M} \sum_{j=0}^{M-1} y(n-j) \tag{4.2}$$

where M is the length of the smoothing window. Setting $M = 3$ provided sufficient smoothing without eliminating (possibly) relevant peaks.

The resulting $z(n)$, shown in the lower graph of Figure 20, has a waveform with clear peaks centred about the QRS-complexes. However, it was observed that $z(n)$ may, in some cases, still contain artefacts with small amplitudes. Therefore, all values in $z(n)$ smaller than the root mean square (RMS) of $z(n)$, calculated with Equation 4.3, are disregarded.

$$rms(z(n)) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} z(i)^2} \tag{4.3}$$

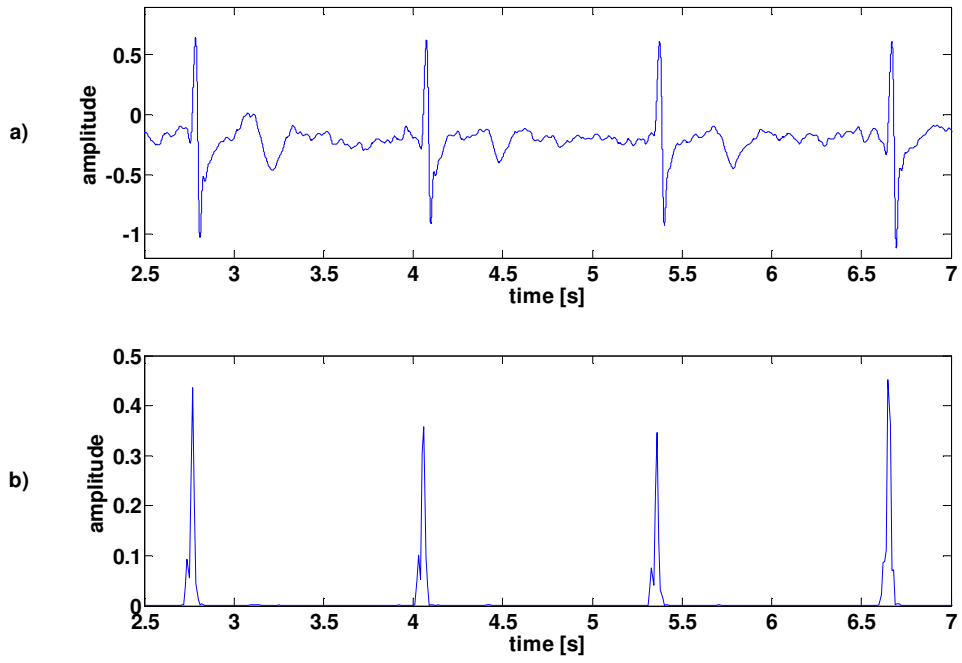

Figure 20: QRS-peak detection: a) ECG recording. b) $z(n)$.

The next step is to identify the location of the peaks in $z(n)$. This is done with the *findpeaks* function built into Matlab, which identifies local peaks in a signal. The time histories of successive identified peaks are used to eliminate false peaks in $z(n)$. False peaks are assumed to occur within 150 ms of the peaks caused by the genuine QRS-complexes. Therefore, all the local peaks identified within a 150

ms window are sorted and every peak except the single local maximum in that window is discarded.

The final step in the segmentation process is to calculate an approximate start of diastole in all the recorded cardiac cycles. Burke and Nasor (2002) developed second order approximation equations for calculating the timing relations of the constituents of the human electrocardiogram. These equations describe the timing of events in the ECG as functions of the time between successive QRS peaks and has been used by previous authors to automate the segmentation of heart sounds (de Vos, 2004), (Visagie, 2007). Equations 4.4 and 4.5 were derived from the ECG's of male subjects, but are used in this study for the female participants as well. Equation 4.6 was adopted to more accurately calculate the start of diastole in the ECG signals.

$$T_{Q-T} = 1.65T_{R-R}^{1/2} - 0.84T_{R-R} - 0.46 \tag{4.4}$$

$$T_{T-wave} = 1.29T_{R-R}^{1/2} - 0.66T_{R-R} - 0.42 \tag{4.5}$$

The estimated onset of diastole in each cardiac cycle is then calculated as:

$$T_{Diastole\,start} = \frac{T_{Q-T} + T_{T-wave}}{2} \tag{4.6}$$

where T_{R-R} is the R-R time interval (in seconds) between successive QRS peaks. $T_{Diastole\,start}$ is the estimated time after the start of the QRS complex, at which diastole begins. Although equations 6 and 7 were derived from normal hearts only, it was found to be effective and provided accurate results for all the recordings made in this study, including those of abnormal hearts. This was established through a visual inspection of the segmentation process on all the recordings.

The result of the segmentation algorithm applied to a normal heart sound is illustrated in Figure 21. It shows how the heart sound cycles are successfully divided into the systole and diastole parts, even though the ECG signal exhibits a significant trend. The segmentation of an abnormal heart sound is shown in Figure 22. The patient presented a very weak ECG signal corrupted by spurious valleys and crests. The performance of the segmentation algorithm can be verified by comparing the synchronised ECG trace with the heart sound in Figure 22a).

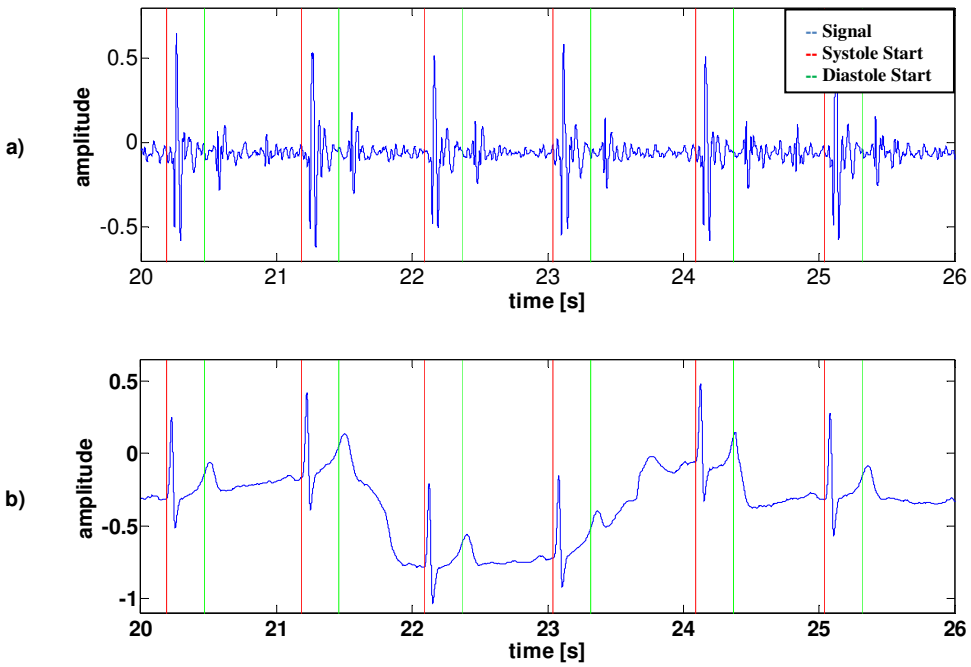

Figure 21: a) Segmented normal heart sound. b) Synchronised ECG trace.

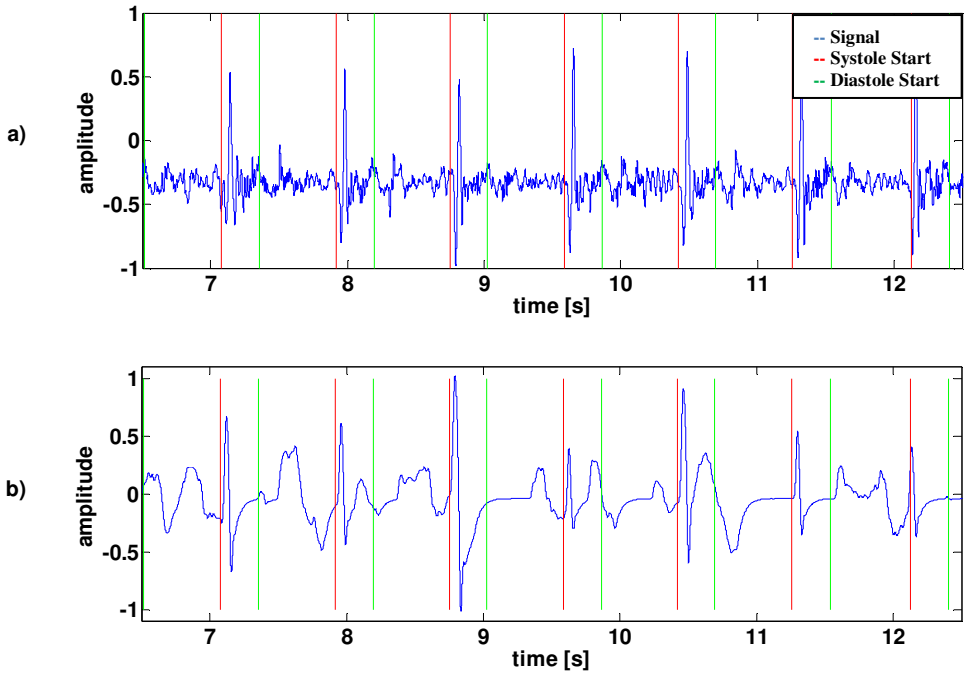

Figure 22: a) Segmented abnormal heart sounds recording. b) Synchronised ECG trace.

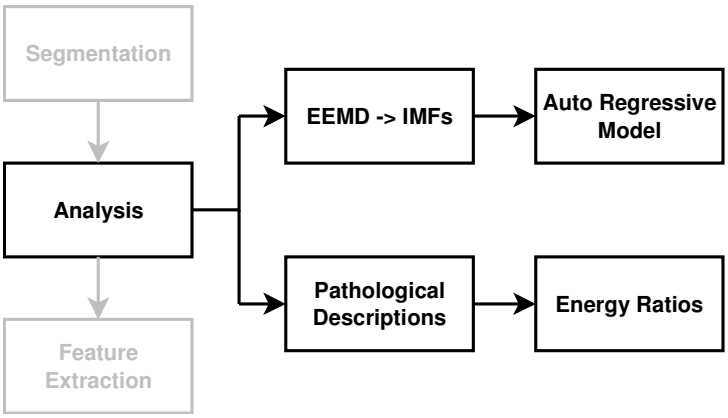

Figure 23: Heart sound analysis flow diagram.

4.4 Analysis

Analysis of the heart sounds is performed to identify the mechanisms involved in producing the various components of the sounds and to facilitate the extraction of features. The analysis, outlined in Figure 23, consists of two components. The first follows a physiological approach, based on discussions with a cardiologist (Lubbe, 2009), who has been trained in cardiac auscultation. The second approach makes use of the ensemble empirical mode decomposition algorithm (EEMD) to analyse the heart sounds and use auto regressive modelling as a quantitative measure for constructing features.

4.4.1 Pathological Approach

Four cases of heart sounds are considered in the pathological approach: Mitral stenosis (MS), mitral regurgitation (MR), aortic stenosis (AS), and aortic regurgitation (AR). These cases are treated separately in this analysis, although cardiac patients may present them in any number of combinations. The physical causes and the abnormal sounds resulting from these cases are discussed in this section and phonocardiograms which illustrate some of these cases are shown.

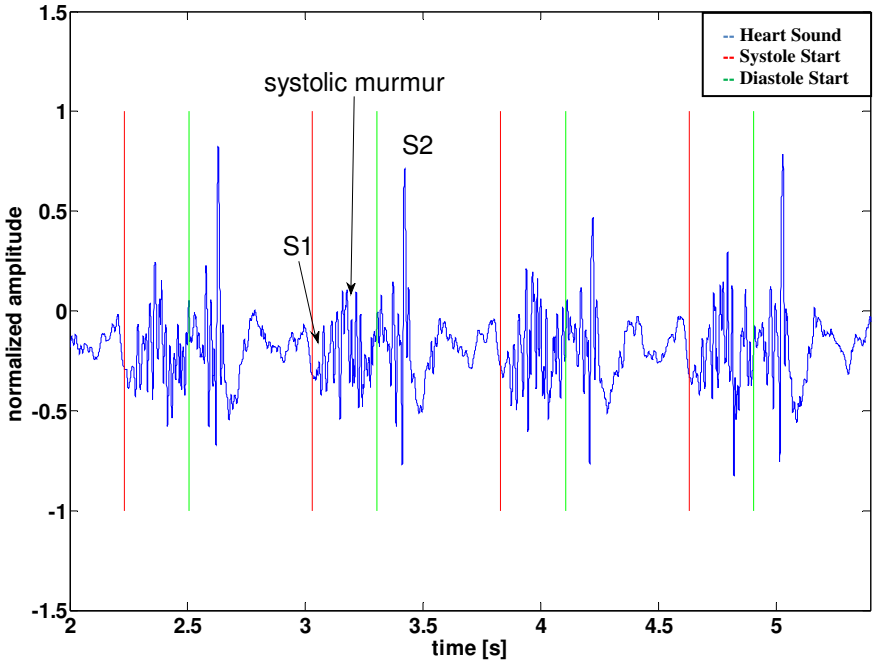

Figure 24: Heart sound of mitral regurgitation showing a soft S1 followed by a pansystolic murmur.

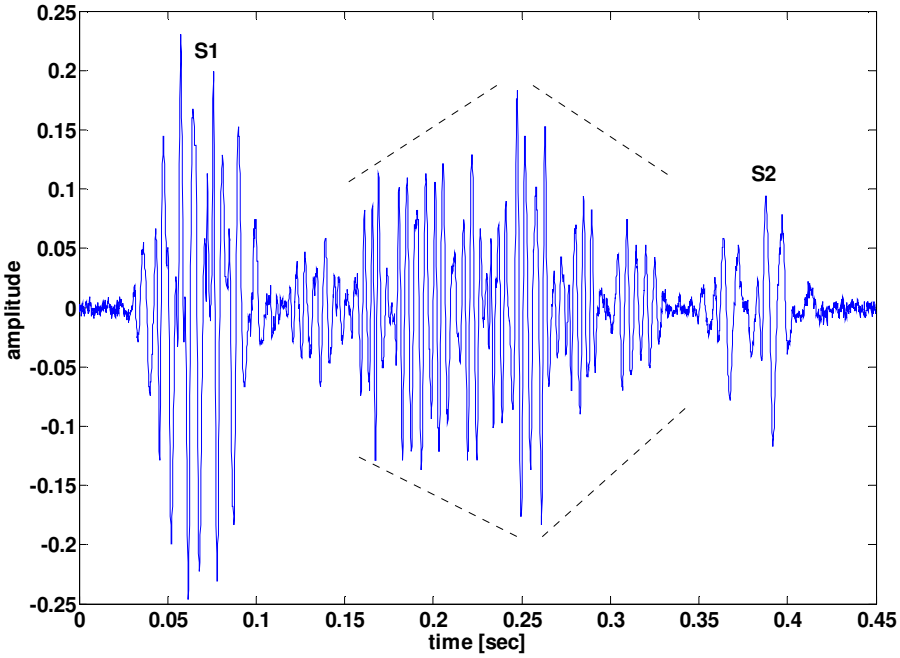

Figure 25: Heart sound recording of Aortic stenosis showing a crescendo-decrescendo systolic murmur (Visagie, 2007).

Mitral Stenosis

Mitral stenosis is a heart valve disorder in which the leaflets of the mitral valve do not open properly, restricting blood flow from the left atria to the left ventricle. This results in an increased pressure gradient of blood across the valve. Common causes of MS include rheumatic fever, calcium deposits around the mitral valve, radiation treatment of the chest and some medications (Lindberg et al., n.d.).

The sounds associated with MS starts with an abnormally loud S1 due to the mitral valve closing with increased force. An opening snap may be heard during S2, due to the forceful opening of the valve. A mid-diastolic, low-pitched, rumble is heard after the opening snap which increases in intensity towards the end of diastole, as the atria contracts (Lubbe, 2009). The rumble is best heard at the apex of the heart with the bell of a stethoscope. The murmur can be accentuated by having the patient roll on his or her left side (Lindberg et al., n.d.).

Mitral Regurgitation

Mitral regurgitation refers to the leaking (backflow) of blood through the mitral valve from the left ventricle to the left atrium. It is a long-term disorder which tends to worsen gradually. MR can be congenital or chronic. Causes of chronic MR include rheumatic heart disease, atherosclerosis, endocarditis, Marfan syndrome, and high blood pressure (Lindberg et al., 2009).

The sound caused by MR is a soft S1 followed by a pansystolic “thrill” murmur (Figure 24). S3 and an S4 gallop may also be heard during auscultation (Lubbe, 2009).

The sounds produced by ventricular septal defect (VSD) and tricuspid regurgitation (TR) are similar to the case of MR and are therefore not treated separately in terms of data analysis. TR is the backflow of blood from the right ventricle into the right atrium and VSD is an opening in the cardiac tissue separating the left and right ventricles. The murmur of VSD radiates from the 4^{th} inter costal (IC) space towards the right, while in the case of MR the murmurs radiate from the 5^{th} IC space towards the left armpit (Lubbe, 2009)

Aortic Stenosis

Aortic Stenosis refers to the aortic valve not opening fully, restricting blood flow from the left ventrical to the aorta. Aortic stenosis can be congenital or chronic. Chronic AS mostly occurs in patients who have had rheumatic fever (Lindberg et al., 2009).

The sound produced by AS is a crescendo-decrescendo ejection systolic mur-

mur (Figure 25). An abnormal click and S4 may also be heard in some cases (Lubbe, 2009).

Aortic Regurgitation

Aortic regurgitation is the backflow of blood into the left ventricle after it was pumped into the systemic circuit. The causes of AR include congenital heart disease, endocarditis and rheumatic fever (Lindberg et al., 2009).

The sound produced by AR is a diastolic murmur called an S3 gallop. It can usually be heard after S2 and decrescendos towards the end of diastole. Pulmonic regurgitation, the backflow of blood into the right ventricle from the pulmonary artery, produces sounds similar to that of AR. Pulmonic regurgitation is therefore treated together with AR in terms of the data analysis (Lubbe, 2009).

4.4.2 Ensemble Empirical Mode Decomposition Analysis

For analysing nonlinear and non-stationary data, Haung et al. (1998) developed a technique known as the Hilbert Haung transform (HHT). At the core of the HHT is a method called empirical mode decomposition (EMD) which was originally developed to study ocean currents (Huang et al., 1998). The EMD uses a process called sifting to decompose a signal into a set of time-energy functions called intrinsic mode functions (IMFs). From the IMFs, the instantaneous frequencies within the original signal can be derived (as functions of time) through the use of the HHT. The resulting time-energy-frequency distribution is called the Hilbert Spectrum. Each IMF represents a single oscillatory mode of the original signal, which can assist in the identification of structures and the analysis of local events embedded in the original signal (Ari & Saha, 2008).

The sifting process that was used to extract the IMFs from a heart sound recording can be described as follows:

i. $X(t)$ represents a single heart sound cycle.

ii. Initialize, $i = 1; r(i) = X(t)$.

iii. Identify local extrema of $r(i)$ and calculate the upper envelope, e_U, and lower envelope, e_L, using cubic spline interpolation.

iv. First sifting component, $h = r_i - (\frac{e_U + e_L}{2})$.

v. Redo steps iii & iv, treating h as the new $r(i)$, until stopping criteria are satisfied.

vi. The first IMF is then $c_1 = n$.

vii. Separate c_1 from the data, $r(i+1) = r(i) - c_1$.

viii. $i = i + 1$.

ix. Extract IMF components $c_1, c_2 \ldots c_n$ by repeating from step iii until c_n is a monotonic function and no more IMFs can be extracted.

Figure 26 shows the upper and lower envelopes (step iii) computed for S1 of a normal heart sound. The mean of these envelopes is used to compute the IMFs in steps iv to vi. The criteria for stopping the sifting process in step v can be related to the definition of an IMF. That is (a) the number of extrema and zero crossings must be equal or differ at most by one and (b) the IMF must be symmetrical with respect to a zero mean (Huang et al., 1998). However, in practice these conditions can lead to over sifting, eliminating physically meaningful information from the IMF. Haung et al. (1998) therefore proposed a technique in which the standard deviation of the current and previous sifting components is evaluated and if found to be below a threshold of 0.3, the sifting is stopped and the previous h is taken as the new IMF.

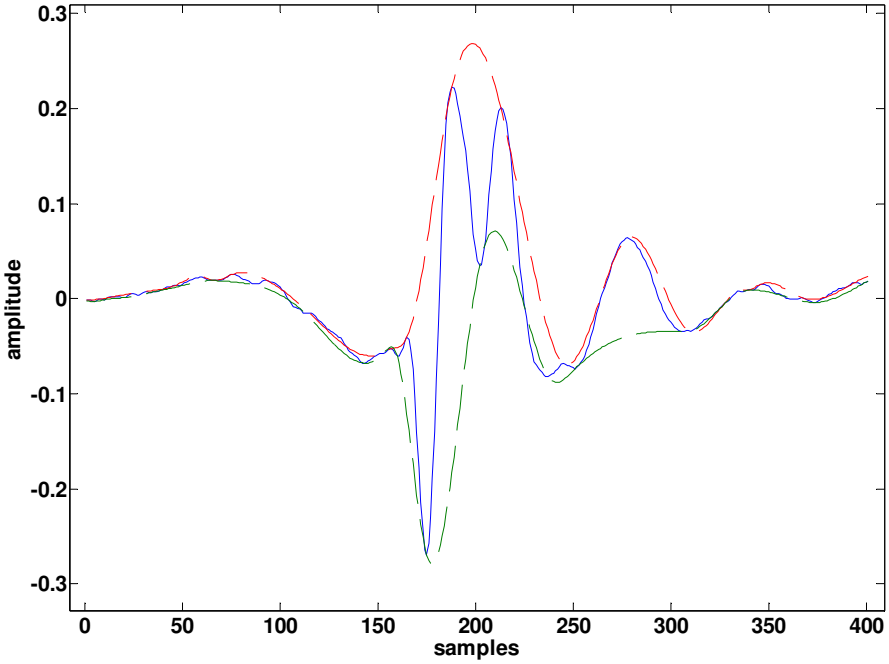

Figure 26: Computation of the intrinsic mode function: the upper and lower envelopes of the local maxima and minima of S1.

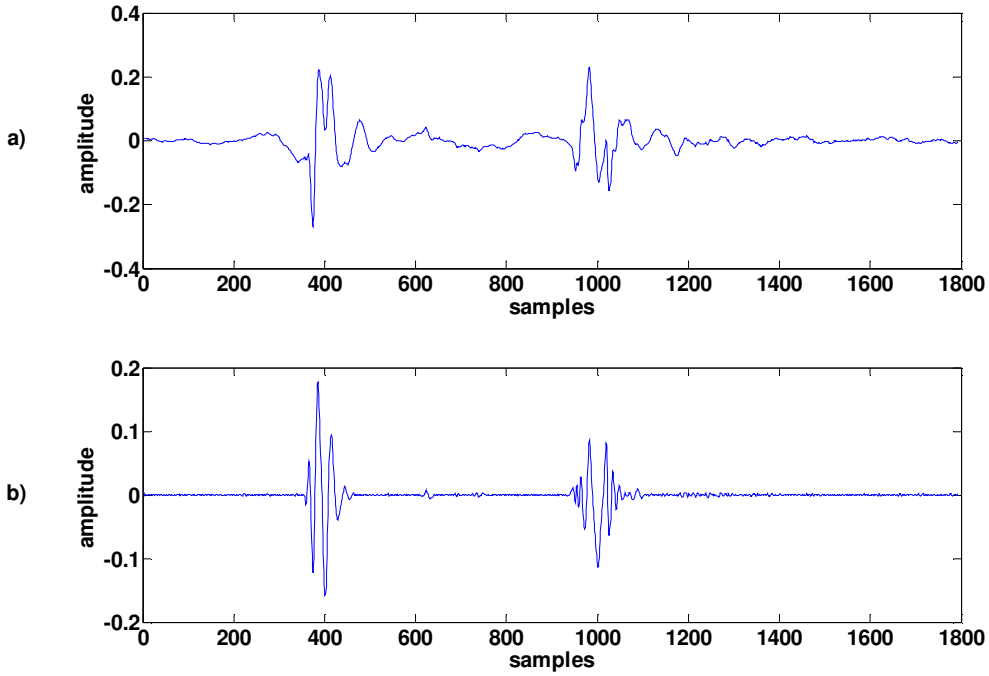

Figure 27: a) PCG device recording of a single heart sound cycle. b) The first intrinsic mode function.

By summing all the IMF components, the original signal can be reproduced:

$$X(t) = \sum_{i=1}^{n} c_i \tag{4.7}$$

The first IMF component, c_1, of a single heart sound cycle is shown in Figure 27b) and the sound from which it was derived, in Figure 27a).

The important properties of the EMD method are:

1. EMD is an adaptive data analysis method that is based on the characteristics of the data alone.
2. EMD makes no *a priori* assumptions about the data.
3. EMD acts as a dyadic filter bank when applied to a series of white (Gaussian) noise (Flandrin et al., 2004).

A major disadvantage of EMD, however, is the frequent appearance of mode mixing (Wu & Haung, 2009). Mode mixing occurs when dissimilar scales of a signal appear in the same IMF or when similar scales are spread over various IMFs. Therefore, to address this problem, an important improvement to the EMD was proposed by Wu and Haung (2009) by transforming it into a noise-assisted data analysis method called ensemble empirical mode decomposition (EEMD). EEMD takes advantage of the characteristics of white noise to provide a frame of reference for the original EMD. Finite amplitude white noise is added to the signal to uniformly fill its whole time-frequency space. Due to the filter bank properties of EMD, the signal components are subsequently projected on the proper scales of reference established by the white background noise.

The EEMD creates an ensemble of white noise added signals by adding a different series of white noise numerous times to the original signal. The i^{th} member of the ensemble would then be:

$$x_i(t) = x(t) + w_i(t) \tag{4.8}$$

where $x(t)$ is the original recorded signal and $w_i(t)$ is the i^{th} series of added white noise. In the mean of a sufficiently large ensemble, the white noise cancels out, leaving only the persistent part of the signal. The outcome of this approach is that the components of similar scales are more likely to gather in the same IMFs, reducing the effects of mode mixing (Wu & Haung, 2009).

The EEMD modification to the original EMD can be outlined as follows:

i. Add a white noise series with an amplitude equal to the ratio, *Nstd*, times the standard deviation of the original heart sound signal to obtain $X_i(t)$.

ii. Decompose $X_i(t)$ into a set of n IMFs, $c_{i,1}, c_{i,2}, \ldots, c_{i,n}$ using EMD.

iii. Repeat i and ii, each time with a different white noise series, to obtain $i = NE$ sets of IMFs, where NE is a sufficiently large number.

iv. Calculate the mean of the ensemble of IMFs to obtain the final result.

The values for $Nstd$ and NE were chosen as 0.1 and 100 respectively. These values are based on experimentation and the implementations of the method by Wu and Haung (2009). Generally, the larger the ratio of amplitude of the added white noise, $Nstd$, the more ensembles members, NE, is required, but at increased computational cost. However, making $Nstd$ too small may not provide the alterations in the signal that the EMD depend upon to properly perform the decomposition.

The EEMD analysis applied to a normal heart sound is shown in Figure 28. The EEMD was able to successfully capture the first heart sound, S1, in the first IMF component with the second heart sound initially appearing only in the second IMF. Comparatively, the EEMD analysis performed on an abnormal heart sound with diastolic murmurs and lower pitched systolic murmurs is shown in Figure 29. In this figure, the diastolic and systolic mummers are encircled in the IMF components. Especially the diastolic murmurs are clearly distinguishable from the first and second heart sounds in these IMFs. These figures illustrate the usefulness of the EEMD method, particularly for the purpose of detecting abnormalities in heart sounds.

The IMFs containing the most physically relevant and meaningful components of the heart sound recording is not necessarily the first or second. In some of the recordings the first IMF would only contain small amplitude, high frequency noise components which are irrelevant to the analysis. Therefore, a method has been devised with which to automatically judge the relevance of IMF components.

The method involves calculating the correlation coefficient, ρ, between the heart sound and each IMF:

$$\rho_{sound,imf} = \frac{COV(sound, imf)}{\sigma_{sound}\sigma_{imf}} \tag{4.9}$$

where the nominator is the covariance between the heart sound and the IMF in question and the denominator contains the standard deviations of the heart sound and IMF. The three IMFs which yield the highest correlation to the heart sound from which it was decomposed, is then chosen as significant for further analysis while the others are discarded. Determining which IMFs are significant is important as it significantly reduces the amount of irrelevant information input to the classification system.

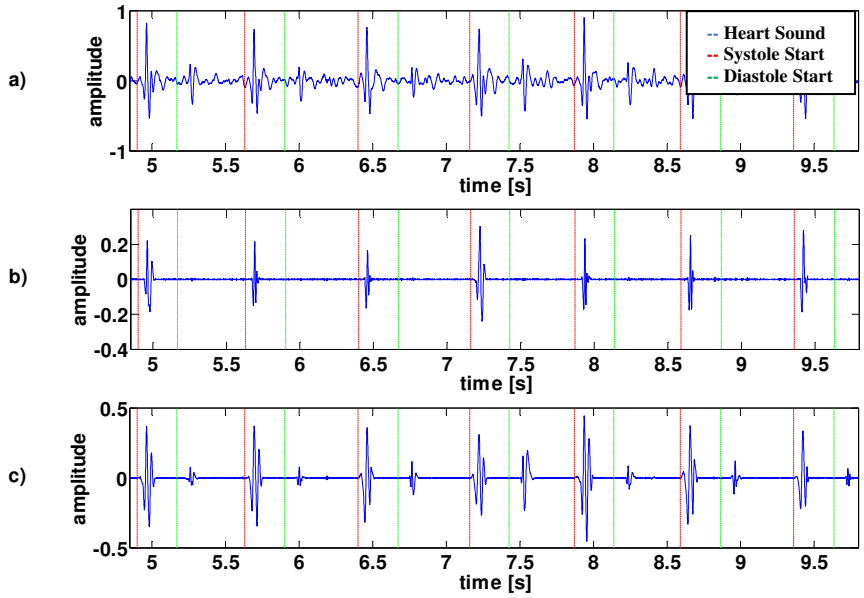

Figure 28: The EEMD method applied to a normal heart sound: a) Normal heart sound. b) The first IMF component. c) The second IMF component.

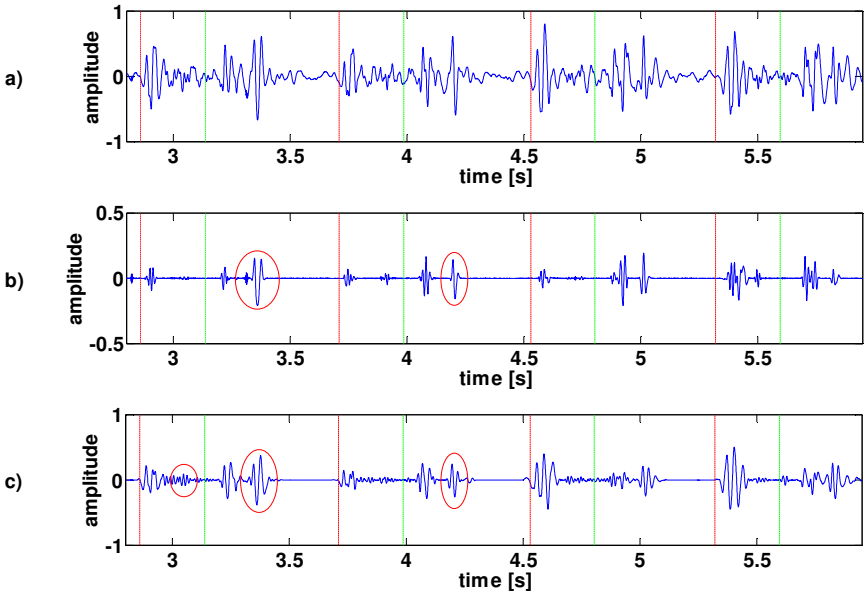

Figure 29: The EEMD method applied to an abnormal heart sound: a) Abnormal heart sound with systolic and diastolic murmurs. b) The first IMF component. c) The second IMF component.

4.4.3 Auto Regressive Model

In order to characterise the spectral information in the intrinsic mode functions extracted with the EEMD analysis, autoregressive modelling was performed. An auto regressive model was computed for every systole and diastole cycle of each of the three IMFs which has been identified to contain significant information. The performance of the modelling was judged based on how well the classification scheme separated the normal and abnormal heart sound classes.

Auto regressive modelling forms part of the parametric group of methods for estimating the power spectra of linear systems. These methods are designed for use where only short data records are available, which makes them ideal for characterising IMFs. Advantages of auto regressive models include the ease of calculating model parameters and its ability to represent spectra with narrow peaks, or resonances (Proakis & Manolakis, 2007).

The auto regressive model parameters were estimated using the unconstrained least squares method as is given by Proakis & Manolakis (2007). This method was chosen as its performance characteristics have been found to be superior to most other methods and it is insensitive to problems related to frequency bias and spurious peaks (Proakis & Manolakis, 2007).

The auto regressive model can be defined as:

$$X_t = \sum_{i=1}^{p} \varphi_i X_{t-i} \tag{4.10}$$

where X_t is the signal to be modelled, $\varphi_1, \ldots, \varphi_p$ are the model parameters, and p is the model order. Selecting the model order, p , is an important part of auto regressive modelling. Proakis & Manolakis (2007) suggests that as a general rule if p is selected too low a highly smoothed spectrum is obtained. While, if p is selected too high false low-level peaks may appear in the spectrum. Model orders of $p = 5, 6, \ldots 20$ were evaluated by using the model parameters as features. The classification performance was then computed for each model using cross-validation and the leave-one-out method, as discussed in Section 4.6.4. It was found that using $p = 12$ resulted in superior classification performance compared to other models and was therefore used in the final implementation.

4.5 Feature Extraction

The features extracted from each recorded cardiac cycle ultimately serve as input to the statistical classification system which detects if cardiovascular disease is present. The choice and design of these features therefore have a direct influ-

ence on the performance of the classification process. Through the use of the analysis methods presented in previous chapters, the features are designed to contain the elemental characteristics of the recorded heart sounds.

The features in the final implementation of the autonomous auscultation system amounted to 7 pathology-based time domain features and 72 features based on EEMD. The latter consist of 12 auto regressive model coefficients computed separately for systole and diastole for each of the three significant IMFs.

A total of 79 features per participant were therefore used as input to train and simulate the classification scheme.

4.5.1 Time domain, pathology-based Features

Energy ratios of different parts of the heart sound cycle have been identified for providing information relevant to the detection of CVD. The choice of energy ratios is based on discussions with a cardiologist, who has training in cardiac auscultation (Lubbe, 2009), and follows from the pathological analysis of Section 4.4.1. Once the energy ratios are calculated they are each used as input to the classification scheme.

Feature extraction starts by dividing each systole and diastole part of each heart sound cycle into three equal pieces as illustrated in Figure 30 and Figure 31. The average energy, calculated with Equation 4.3, contained in each of the sections is labelled as either $ES_{1,\ldots,3}$ for systole or $ED_{1,\ldots,3}$ for diastole. The identified energy ratios are listed in the equations shown in Table 4.

Table 4: Pathology-based, time domain features.

Systolic Feature	Diastolic Feature	Comparison of Systolic and Diastolic regions
$F_1 = \frac{ES_1}{ES_2 + ES_3}$	$F_3 = \frac{ED_1}{ED_2}$	$F_2 = \frac{ED_2 + ED_3}{ES_1}$
$F_7 = \frac{ES_1}{ES_2}$		$F_4 = \frac{ES_1}{ED_1}$
		$F_5 = \frac{ES_2 + ES_3}{ED_2 + ED_3}$
		$F_6 = \frac{ES_1 + ES_2 + ES_3}{ED_3}$

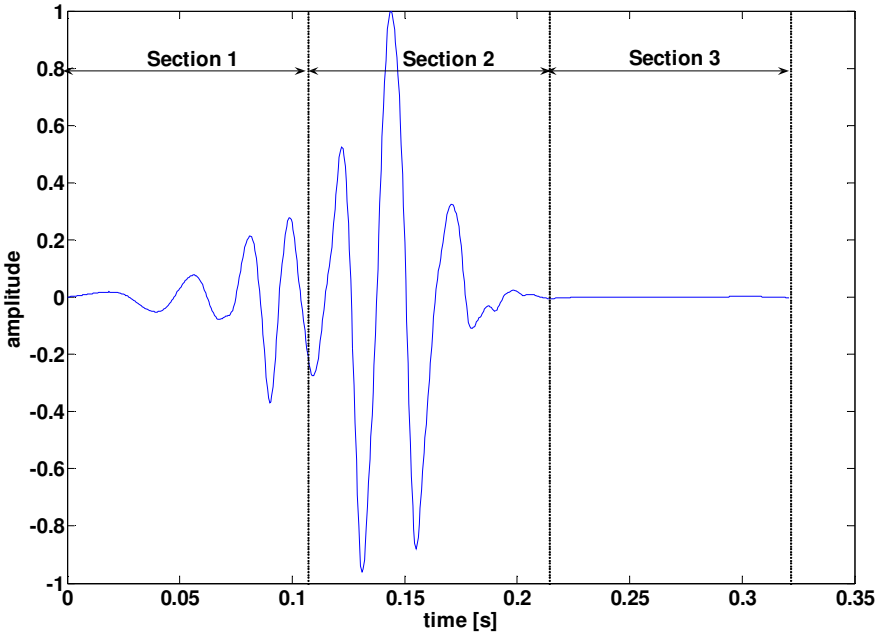

Figure 30: Division of systole heart sound into three sections.

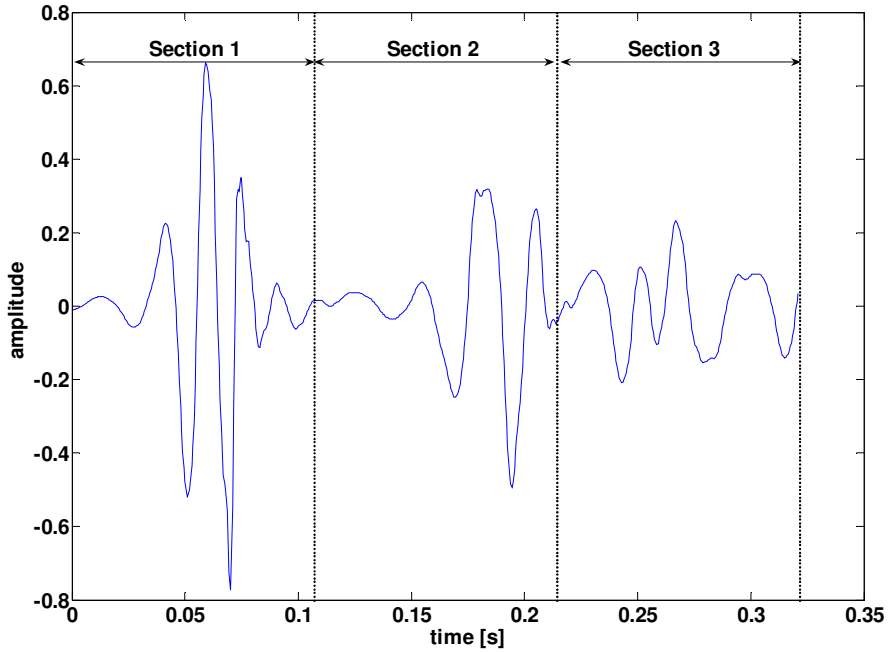

Figure 31: Division of diastole heart sound with murmurs into three sections.

4.5.2 Auto Regressive Model Features

The 12^{th} order auto regressive model parameters described in Section 4.4.3 are calculated for each systole and diastole part of the each recorded cardiac cycle. This is done for all three of the significant IMFs derived with the methods of Section 0. The result is a total of 72 features for every recorded cardiac cycle for every participant.

4.5.3 Feature Averaging over Cardiac Cycles

Utilizing each recorded cardiac cycle's features as a separate data point input to the classification scheme resulted in poor classification performance and subsequent separation between of the normal and abnormal classes. This is thought to be caused by the extreme cycle-to-cycle variation exhibited by the heart sound components and the fact that an unequal amount of cardiac cycles are available for each of the participants.

Therefore, the statistical median is calculated over the features of all the recorded cardiac cycles of each participant. The median, η, which minimizes the average and absolute deviations of a dataset, is considered a more representative measure of the locality of a dataset with a skew probability distribution than the mean, μ, which is generally used(Steyn et al., 2004). The median is defined statistically as the value, η, for which the probability is:

$$P(X \leq \eta) = 0.5 \tag{4.11}$$

X, in this case, is the vector containing the values of a particular feature of the all recorded cardiac cycles of a single participant. Averaging over the cardiac cycles to produce a single data point per participant was found to generally yield better classification results than using each cycle as a separate data point.

4.6 Classification using Artificial Neural Networks

Classification is the task of using computer algorithms to autonomously find patterns in data, which is a fundamental challenge. Its main concern is the action of organizing data into different categories based on those patterns. A well-known application of this process is handwriting recognition, found on many handheld computers, where a handwritten digit is recognised and its identity is produced $0,\ldots,9$ as the output (Bishop, 2006).

Given an input vector $\mathbf{x}$ the goal of classification is to assign it to one of K

discrete classes, C_k, where $k = 1,\ldots,K$. The classes are generally disjoint so that each input can be assigned to only one class (Bishop, 2006). In this study, each recorded participant is represented by a single, unique input vector, $\mathbf{x}$, whose elements are the features described in the previous sections. The number of elements in $\mathbf{x}$ determines the dimensionality, D, of the input space.

The target variable t is the vector we wish to predict and is a binary representation of the class to which each individual observation, $\mathbf{x}$, is mapped. In the case of a two-class problem, such as implemented in this study,

$$t \in \begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix} \tag{4.12}$$

where $\mathbf{t} = \begin{bmatrix} 1 & 0 \end{bmatrix}$ represents class C_1 (normal, no CVD) and $\mathbf{t} = \begin{bmatrix} 0 & 1 \end{bmatrix}$ represents class C_2 (CVD is present).

The process of classification is similar to that of fitting a curve to data. The input data is fit to a combination of functions through training to produce a desired output. The classifier divides the feature space of the problem into decision regions whose boundaries are called decision boundaries forming $D-1$ dimensional hyperplanes within the D-dimensional input space. The profile of these boundaries is determined during training, whereby an unknown data point is classified, depending on its location in the feature space, relative to the decision boundaries.

Using linear combinations of fixed basis functions for classification has advantages of useful analytical and computational properties, but presents a significant limitation to their application on large scale problems. This limitation arises due to what is commonly referred to as the curse of dimensionality. As an example, consider the task of polynomial curve fitting. To fit a 3rd order polynomial to D input variables would require D^3 independent coefficients. As D increases, the number of coefficients increase proportional to D^3. Similarly, fitting a polynomial of order M, the number of coefficients would increase by D^M. This example illustrates how the practical application of linear classifiers is limited to small scale problems due to this curse of dimensionality (Bishop, 2006).

To overcome the challenges arising from the high dimensionality of large scale problems, it is necessary to adapt the basis functions to the data. An attractive approach is to fix the number of basis functions in advance, but allow their parameters to be adaptive during training. According to Bishop (2006) the most successful model of this type in pattern recognition is the feed-forward neural network, and is therefore implemented in this study.

4.6.1 Network Construction

The classification scheme implemented in this study is an ensemble of three-layer perceptron feed-forward neural networks. Each neural network consist of interconnecting groups of functions called neurons, which process information and then passes it to the other neurons connected to it. The neurons are grouped into different layers of which the first is the input layer and consists of the features for each participant. The second layer is a hidden since its inputs and outputs are not used directly. The third layer is the output layer and consists of 2 neurons, one for each class. Its output is the final decision output of the network.

During experimentation with different features it was found that fixing the number of neurons in the hidden layer to 1.3 times the number of features generally produced good cross-validation results through the leave-one-out method described in Section 4.6.4. This ratio was also used in the final implementation of the system to determine the number of neurons in the hidden layer.

The overall network function governing the output, y_k, of each ensemble member is given by Equation 4.13 and is illustrated graphically in Figure 32.

$$y_k(\mathbf{x},\mathbf{b},\mathbf{w}) = f\left(\sum_{j=1}^{M} w_{kj}^{(2)} h\left(\sum_{i=1}^{D} w_{ji}^{(1)} x_i + b_j^{(1)}\right) + b_k^{(2)}\right) \tag{4.13}$$

The number of outputs for each network is given by $k = 1,\ldots,K$ in Equation 4.13, where $K = 2$ represents the number of classes to which the input data, $\mathbf{x}$, will be classified. The number of neurons, M, in the hidden layer is given by $j = 1,\ldots,M$. The dimensionality of the input data is given by $i = 1,\ldots,D$, where D represents the number of features for each participant. The functions f and h are the activation functions of the neurons. The output layer of each network is unconstrained with the linear activation function, f, given by Equation 4.15. The activation function, h, of the neurons in the hidden layer is the hyperbolic tangent sigmoid function of the form given by Equation 4.14.

$$h(x) = \frac{1}{1+\exp(-x)} \tag{4.14}$$

$$f(x) = x \tag{4.15}$$

The adaptable parameters of each network in the ensemble are the weights and biases. For the neurons in the output layer, these parameters are represented by $w_{kj}^{(2)}$ and $b_k^{(2)}$ in Figure 32 and similarly for the neurons in the hidden layer they are represented by $w_{ji}^{(1)}$ and $b_j^{(1)}$. The construction of each network in the final ensemble implementation consisted of $M = 103$ hidden neurons and $D = 79$ dimensions in the feature space.

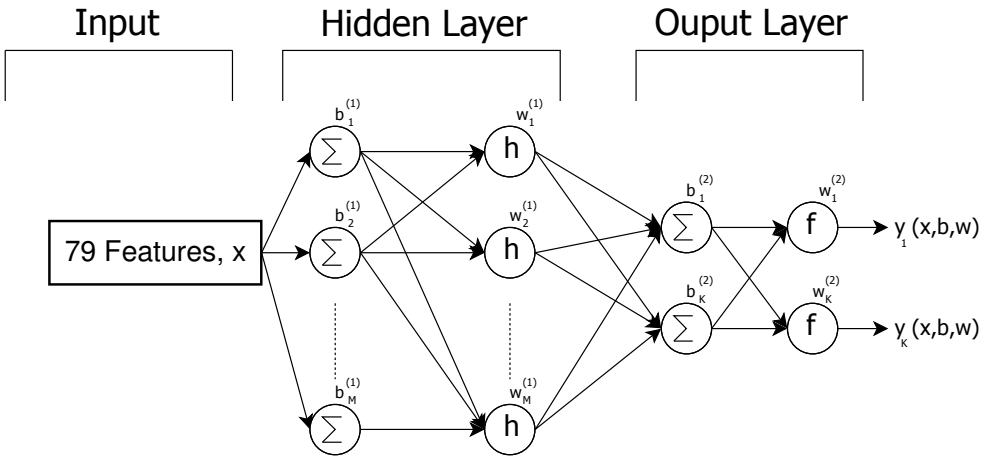

Figure 32: Artificial neural network graphical representation.

Each ensemble neural network member is initialised before it is trained using the Nguyen-Widrow method (Nguyen & Widrow, 1990). This method generates initial weights and biases in such a way that the active regions of each layers' neurons are distributed more or less evenly over the input space. The advantage of this method over a solely random initialisation is that training is performed considerably faster.

4.6.2 Ensemble Neural Networks

Ensemble neural networks have been applied in many fields and have been shown to generally give more accurate results compared to unitary networks(Bhatikar et al., 2004). Ensembles are typically used when the data is too complex to be modelled by a single network. The objective of the ensemble is to combine the individual outputs of each neural network member to obtain better generalisation.

In this study an ensemble of 99 neural network members is implemented to obtain a final decision output and confidence. The number of members was chosen as large as possible without severely affecting computation performance. A choice of 100 members adequately fulfilled this requirement, but was reduced to the closest odd number to prevent decision ties. Liu et al. (2004) implemented ensemble artificial neural networks in their study of classifying gene expressions and found that it repeatedly provided a better accuracy than a single network. Bhatikar et al. (2004) implemented a decision-of-experts ensemble ANN classification scheme in their cardiovascular screening system for paediatrics and reported a sensitivity and specificity of 88% and 83% respectively.

The ensemble decisions are combined based on the number of times a particular participant was classified to a certain class. For example in Table A1, participant 3 was classified by 87 ensemble members as normal and only by 12 as abnormal. The final output is thus normal with a confidence of 87.9%. The confidence is calculated as a percentage with the formula,

$$confidence = \frac{\#\,decisions\ in\ output\ class}{\#\,ensemble\ members} \tag{4.16}$$

Simulating each participant with multiple ensemble members was necessary to obtain a measure of how well the features separated him or her into a particular class. Simulating the masked participant only once could produce an unrealistic result as he or she might have been classified correctly (or incorrectly) only by chance. Therefore, re-simulating 99 times provided a decision confidence on the classification accuracy of the entire system on each participant.

4.6.3 Training

Supervised training is applied to the network and involves adapting the weights and biases in order to improve the match between the network output and the target values assigned for each of the input vectors.

The backpropagation algorithm (Bishop, 2006), which sends information backward and forward through the network, is used to calculate error gradient information. Using these gradients, the weights and biases are adapted with the scaled conjugate gradient descent optimization algorithm to minimize the squared differences between the outputs and targets.

Weight and bias learning was performed with the gradient descent with momentum algorithm (Vanderplaats, 2005). The learning rate was set to 0.01 and the momentum term to 0.9. A training goal of 1×10^{-5} was used. Network performance was evaluated during training using the mean-squared-error performance function.

The training data presented to each ensemble neural network member was divided into two parts: 90% train and 10% validate. The 10% validation is used during training to automatically fine tune the network parameters and is not related to the cross-validation procedure discussed in Section 4.6.4.

4.6.4 Leave-one-out Cross-Validation

Testing a classification scheme generally yields the likelihood that a future pa-

tient will be recognised and classified correctly. Cross-validation is the process of testing the already trained network with data that has not been used for training and is hence unknown to the network. The cross-validation techniques discussed in this section is not related to the verification and validation of software for mission critical applications. It is, however, necessary to ensure that the ensemble neural network has not overfitted the training data, but sufficiently generalized over the natural variations between individuals.

The customary method for testing a classification scheme, when large quantities of data are available, involves dividing the data into two subsets. The first set is used for training while the second is used as a cross-validation dataset with which to test the entire system. The advantage of this method is that it requires only a few simulation runs, but it is severely dependent on the quantity of available test data. The test dataset also needs to be large enough to be representative of the population from which it is drawn; otherwise the performance result of the network will not be physically realistic. Unfortunately, due to the limited quantity of data available in this study, such testing of the system could not be performed. Therefore, the results from cross-validation are used instead to determine the overall system performance.

Cross-validation is a technique for determining how a classification scheme will generalise over an independent data set. The technique involves partitioning data into subsets, where one subset is used to train the classifier and the other is used as a validation dataset. 10-fold cross-validation is commonly employed whereby the data is divided subsets containing 10 data points each. Training is done on 9 data points and the classifier is tested on the remaining one. The average accuracy obtained over all the subsets is then reported.

The leave-one-out cross-validation method is used to circumvent the constraint caused by the lack of sufficient testing data in this study. The advantage of this method above methods such as 10-fold or 5-fold cross-validation is that almost none of the potential training data is used for cross-validation. However, the method is computationally expensive since the classifier needs to be re-trained and simulated as many times as there are data points (participants in the clinical study). Guo et al. (1994) implemented the leave-one-out method to evaluate the performance of their ANN classifier that was designed to assist in the auscultation of bioprosthetic heart valves (Guo et al., 1994). Bhatikar et al. (2004) developed a screening system for paediatrics based cardiac auscultation and used the leave-one-out method to evaluate its performance (Bhatikar et al., 2004).

The leave-one-out cross-validation algorithm and ensemble implementation is shown in Figure 33. The method starts by masking the data of the first participant, using all the other participants as training data. Then, the first ensemble member is trained and simulated with the masked participant. The inner loop re-initializes the network using the Nguyen-Widrow method, creating the second ensemble

member, which is then re-trained and used to re-simulate the masked participant. The inner loop now iterates until the masked participant has been simulated with all the ensemble members. The second participant is now masked and a new training dataset is formed. This process is repeated until every participant has been simulated with all the ensemble members.

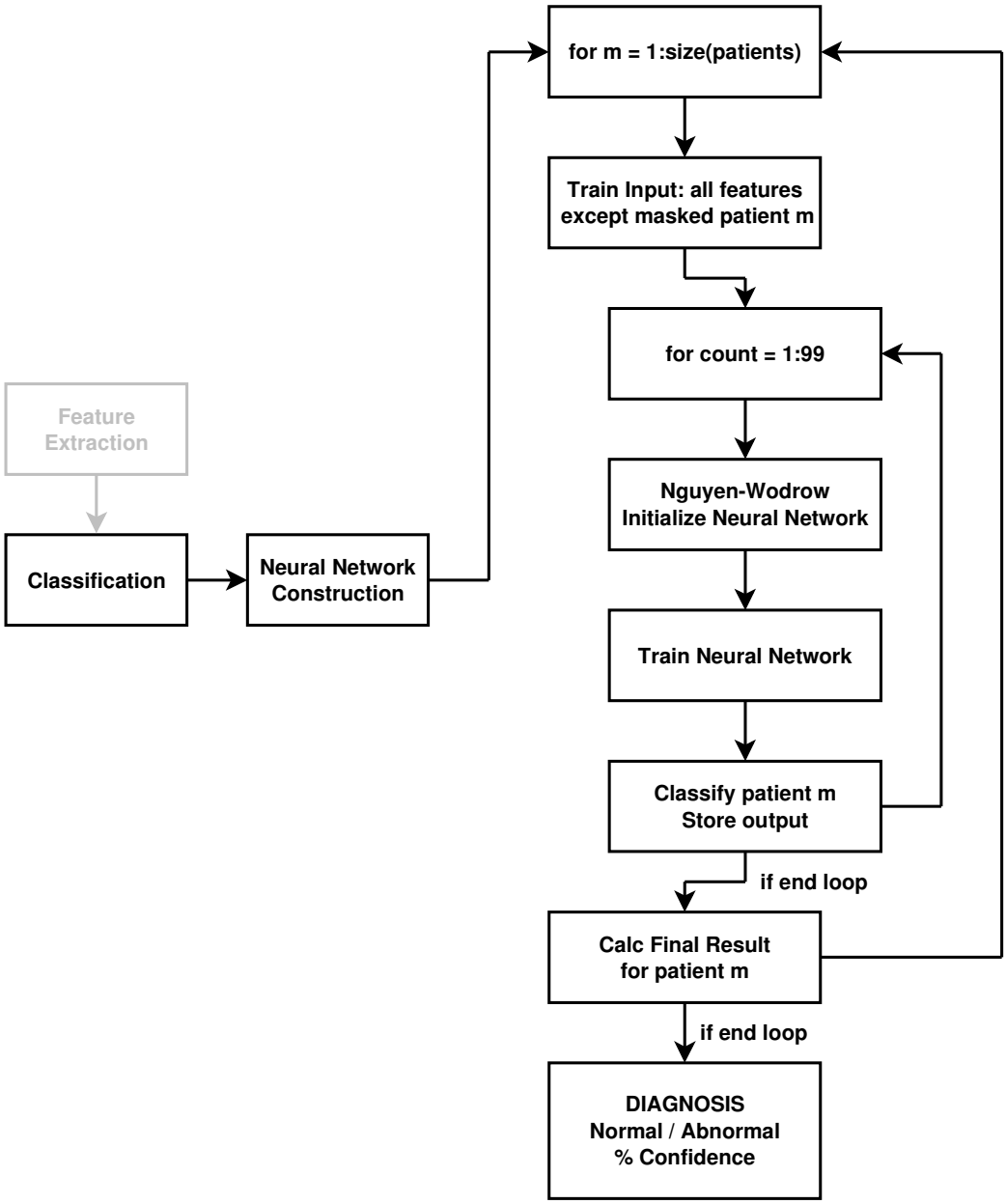

Figure 33: The leave-one-out cross-validation and ensemble neural network algorithm.

4.7 Graphical User Interface

A graphical user interface (GUI) was created as part of the development of the autonomous auscultation system. It was designed to enable fast and easy verification of the functioning of all the sub-algorithms on any of the participants. Figure 34 show how the GUI is used to verify that the segmentation algorithm correctly segmented the heart sound of a particular participant. Figure 35 illustrates how the GUI assists with the analysis of the heart sounds by displaying the result of the EEMD algorithm on the heart sounds.

The GUI can also be used for displaying phonocardiogram information for the purpose of providing decision support to clinicians. Further development of the GUI could even provide the potential for it to be incorporated into current electronic medical record systems in healthcare facilities to graphically portray patient information.

Figure 34: Autonomous auscultation system graphical user interface showing the heart sound segmentation process.

Figure 35: Autonomous auscultation system graphical user interface showing the ensemble empirical mode decomposition analysis of a heart sound.

4.8 Chapter Summary

This chapter discusses the detail design and implementation of the components constituting the autonomous auscultation software system. These components include importing data, denoising, segmentation, feature extraction and classification. Evaluation of the system through the leave-one-out cross-validation method is also described. The chapter concludes with a description of the graphical user interface which has been designed to visually verify the functioning of all the subcomponents of the system.

5. RESULTS AND FINDINGS

This chapter discuss the performance results of the autonomous auscultation system. Through implementation of the leave-one-out cross-validation method discussed in Section 4.6.4 a decision output was obtained for every participant. Combining of the ensemble neural network members' outputs to obtain a final decision is illustrated and is related to the echocardiograph used as the benchmark diagnosis (gold standard). A description of the overall performance of the system is presented through sensitivity, specificity, accuracy, predictive values and likelihood ratios. Finally, a comparison with the results of other, similar, studies is drawn.

5.1 Individual Participant Results

The leave-one-out cross-validation results were used during development of the autonomous auscultation system to test its performance as well as to verify the functioning of all its subcomponents. If any of the processes prior to classification malfunctioned, it was expected to be reflected in these outputs or at worst by the decision confidences. This assumption is founded on the fact that the network is cross-validated with data which has not been used for training.

The final results of the autonomous auscultation system on participants recorded with the PCG device are given in Table A1 and similarly, those recorded with the Auscultation Jacket are listed in Table A2. These tables summarise the result of the echocardiograph on each participant, the final cross-validation output of the autonomous auscultation system and a confidence percentage on the decisions. An extract from Table A1 is shown in Table 5 to illustrate the results. For example, participant #56, in Table 5, presented an abnormal heart condition which was diagnosed as aortic stenosis (AS) by the echocardiograph (2^{nd} and 3^{rd} columns). The number of ensemble members that classified each participant as either class 1 (normal) or class 2 (abnormal) is listed in the 4^{th} and 5^{th} columns. Participant #56 was classified by 12 members as normal and by 87 as abnormal. This participant is therefore identified as abnormal with a final classification output of $[0 \quad 1]$ and a confidence of 87.9%. The confidence is calculated with Equation 4.16 as a percentage based on the number of neural network members that correctly classified the participant. If the confidence value is within ± 5% of 50%, then the diagnosis can be regarded as inconclusive.

Table 5: Extract from cross-validation study.

	Echocardiograph		Ensemble decisions		Classification output			
Participant	Abn/Norm	Diagnosis	Class 1	Class 2	Class 1	Class 2	Abn/Norm	Confidence
56	A	AS	12	87	0	1	A	87.9%
54	A	prosthetic MV	61	38	1	0	N	61.6%

The incorrectly classified participants are printed in red in Tables A1 and A2. For comparison, the incorrectly classified participant #54 in Table 5 was misclassified by most of the ensemble members largely because the pitch of the clicking sound of his prosthetic mitral valve (MV) falls outside the frequency response of the electronic stethoscopes. In this case, 61 ensemble members falsely identified the participant as normal and only 38 as abnormal. Therefore, the final classification output $[1 \quad 0]$ signifies that the system identified the participant (falsely) as normal.

To illustrate the raw ensemble neural network outputs for each participant, the outputs of participant #35 (Appendix A, Table A2) is listed in Table A3. An extract from Table A3 is shown below in Table 6. The second and third columns contain the outputs for each of the 99 ensemble members. Each member's final decision (4^{th} and 5^{th} columns) and is based on which of these outputs are the largest.

Ideally the output of each member should be close to either $[1 \quad 0]$ or $[0 \quad 1]$, like that of member #2. However, some outputs are problematic such as those of members #1 and #21 in Table 6. In future work, such problematic outputs could be addressed by constructing the network to only have a single neuron in the output layer, or by assigning adaptive weights to the members. If a member then outputs such a problematic result, its influence on the final decision would be lowered by automatically adjusting its weight value.

Table 6: Ensemble neural network output and decisions.

	Ensemble Neural Net Output		Neural Net Decision	
Member	Class 1 (N)	Class 2 (A)	Class 1 (N)	Class 2 (A)
1	0.9677	0.9255	1	0
2	0.0166	0.9974	0	1
...	...	...	...	...
20	0.0183	0.9744	0	1
21	0.9813	0.9633	1	0
22	0.0036	0.9801	0	1
...	...	...	...	...
99	0.0361	0.9910	0	1
		Total Decisions:	11	88

5.2 Autonomous Auscultation System Performance

The performance of the autonomous auscultation system was evaluated separately for the data recorded by the PCG device and the Auscultation Jacket. Due to the hardware differences between the two devices, such as bell and diaphragm materials and stethoscope positioning, the characteristics of their sound recordings were fundamentally different. Consequently, when their recordings were presented collectively to the autonomous auscultation system, the system was unable to differentiate between the normal and abnormal heart sounds. But, when presented independently, differentiation between the normal and abnormal classes were achieved.

The diagnostic value of any medical test, when compared to a gold standard, is conventionally stated in terms of the sensitivity, specificity, predictive values, accuracy and likelihood ratios (Greenhalgh, 1997). Each of these performance measures provide a unique insight into the screening capability of the medical test and is necessary to allow different diagnostic systems to be compared. The notation for expressing these performance measures are listed in Table 7 in the form of a confusion matrix. The rows represent the gold standard and the columns are the cross-validation test results. In the case of this study, the number of true negatives (#TN) is the number of normal, healthy participants classified correctly by the autonomous auscultation system as normal, whereas the number of false negatives (#FN) is the number of abnormal participants classified incorrectly as normal. Similarly, the number of true positives (#TP) is the number of abnormal patients (CVD present) classified correctly by the autonomous auscultation system as abnormal and the number of false positives (#FP) is the number of normal patients classified incorrectly as abnormal.

Table 7: Notation for expressing the result of a validation, screening, or diagnostic test (Greenhalgh, 1997).

normal (gold standard)	#TN	#FP
abnormal (gold standard)	#FN	#TP
	normal (test)	**abnormal (test)**

Confusion matrices of the results from the leave-one-out cross-validation of Section 4.6.4 for the Auscultation Jacket and PCG device are shown in Tables 8 and 9. For example, Table 8 illustrates how the autonomous auscultation system classified 30 healthy participants correctly and 4 incorrectly, whereas 23 abnormal participants were classified correctly and 5 incorrectly.

Table 8: PCG device confusion matrix.

normal (echocardiograph)	30	4
abnormal (echocardiograph)	5	23
	normal	**Abnormal**

Table 9: Auscultation Jacket confusion matrix.

normal (echocardiograph)	28	3
abnormal (echocardiograph)	5	16
	normal	**Abnormal**

The measures used to describe the performance of the developed autonomous auscultation system in this study are summarised in Table 10. Since these values were calculated by comparison with the echocardiograph, they can be used to directly compare this study to other, similar studies.

The sensitivity (Table 10) is the fraction of abnormal participants classified correctly as abnormal. It is a measure of how well the system identified CVD in the heart sounds of the participants in the study. The specificity is the fraction of normal participants classified correctly as normal and is a measure of how well the autonomous auscultation system excluded participants without CVD. The accuracy of the autonomous auscultation system describes the proportion of all the participants that was classified correctly. The sensitivity, specificity and accuracy measures, however, are insensitive to the prevalence of CVD in the test participants. Also, they do not give information regarding the probability that a new, unknown patient will be diagnosed correctly, unless the proportion of abnormal to normal participants in the study is similar to that of the real world (Altman & Bland, 1994). Therefore, predictive values, also referred to as the post-test probabilities, are used as they depend crucially on the prevalence of CVD in the participants being tested.

The positive predictive value is the proportion of participants that tested positive for CVD by the autonomous auscultation system and who are correctly diagnosed. It represents the probability that if a new, unknown, patient is identified as abnormal (CVD present), that he or she really is abnormal. Similarly, the negative predictive value is the proportion of participants that tested negative (CVD not present) by the autonomous auscultation system, when they really are normal.

The usefulness of the autonomous auscultation system as a screening test can further be determined through the use of likelihood ratios computed from the results of the leave-one-out cross-validation study (Greenhalgh, 1997). Likelihood ratios are alternative statistics to the sensitivity and specificity and are similarly used to summarise the diagnostic accuracy of a screening test (Deeks & Altman, 2004).

The positive likelihood ratio gives an indication of how much more likely a positive test is to be found in a future patient with CVD, than in a patient without it. Similarly, the negative likelihood ratio shows how much more likely a negative test result will be found in a patient without CVD, than in an abnormal patient (Greenhalgh, 1997). Generally, the farther the value of a likelihood ratio is from 1, the stronger is the evidence for the presence or absence of CVD. A positive likelihood ratio of above 10 and a negative ratio of below 0.1 are generally regarded as strong enough evidence to rule a diagnosis based only on the test result (Deeks & Altman, 2004).

Knowing how the developed autonomous auscultation system predicts the risk of CVD is essential to clinical practice. The sensitivity, specificity and predictive values cannot be used for this as they are only relevant beyond a study if the participants form a suitable random sample of the population (Deeks & Altman, 2004). The likelihood ratios provide a solution to this as they can be used to calculate the probability of abnormality. For this reason, many clinicians prefer the use of likelihood ratios as performance measures to sensitivities and specificities (Deeks & Altman, 2004). By relating the prior probability and the test probabilities through Bayes' theorem, shown in Equation 4.17, the posterior test probability can be calculated from the likelihood ratios (Bishop, 2006).

$$p(C_k \mid \mathbf{x}) = \frac{p(\mathbf{x} \mid C_k)\, p(C_k)}{p(\mathbf{x})} \tag{4.17}$$

The prior probability, $p(C_k)$, that a patient has CVD, before he or she is tested with the autonomous auscultation system can now be taken into account by multiplying the likelihood ratio with the prior probability. Since the probability of $\mathbf{x}$, $p(\mathbf{x})$, cancels out in Equation 4.17 with a normalisation constant, this multiplication yields the posterior odds, $o(C_k \mid \mathbf{x})$, of a positive or negative diagnosis. The posterior probability, $p(C_k \mid \mathbf{x})$, can then be calculated using Equation 4.18. This posterior indicates the probability that the test patient is indeed positive or negative, given a prior probability and a diagnosis from the autonomous auscultation system (Deeks & Altman, 2004).

$$p(C_k \mid \mathbf{x}) = \frac{1}{(1 - o(C_k \mid \mathbf{x}))} \tag{4.18}$$

Table 10: Performance of the autonomous auscultation system calculated by comparison with the echocardiograph in a cross-validation study.

Performance Measure	Formula	PCG device	Auscultation Jacket
Sensitivity	$\frac{\#TP}{(\#TP+\#FN)}$	82.1%	76.2%
Specificity	$\frac{\#TN}{(\#TN+\#FP)}$	88.2%	90.3%
Accuracy	$\frac{\#TP+\#TN}{(\#TP+\#FN+\#TN+\#FP)}$	85.5%	84.6%
Positive predictive value	$\frac{\#TP}{(\#TP+\#FP)}$	85.2%	84.2%
Negative predictive value	$\frac{\#TN}{(\#TN+\#FN)}$	85.7%	84.8%
Likelihood ratio of a positive test	$\frac{Sensitivity}{(1-Specificity)}$	6.98	7.87
Likelihood ratio of a negative test	$\frac{(1-Sensitivity)}{Specificity}$	0.20	0.26

The sensitivity and specificity results obtained in this study with the autonomous auscultation system are comparable with many similar studies. Bhatikar et al. (2004) reported sensitivity and specificity values of 88% and 83% respectively with their ensemble neural network system to autonomously screen for CVD in paediatrics. Their software system uses recordings made with a standard electronic stethoscope to screen for CVD (Bhatikar et al., 2004). Sinha et al. (2007) reported an accuracy of 95% for their software system which autonomously screens phonocardiograms for pathological conditions (Sinha et al., 2007). Leung et al. (2000) reported sensitivity and specificity values of 97% and 94% respectively in their study of automatic differentiation between normal and pathological phonocardiograms (Leung et al., 2000). Turkoglu et al. (2002) developed a system to autonomously diagnose cardiac valvular disease through auscultation and reported a classification accuracy of 94% in their clinical study (Turkoglu et al., 2002). Chauhan et al. (2008) reported an accuracy of 90% for autonomously detecting diastolic murmurs in heart sounds using their automated screening system (Chauhan et al., 2008).

Unfortunately, references reporting likelihood ratios with regards to the clinical testing of similar autonomous auscultation systems could not be found in litera-

ture. However, the likelihood ratios obtained during this study are sufficiently dispersed from 1 to indicate a high probability the system will identify a future abnormal patient as abnormal and discard a negative patient.

The combination of the results in Table 10 indicates that the developed autonomous auscultation system can be useful as a diagnostic screening tool with high probabilities for producing accurate positive and negative diagnoses.

5.3 Performance Comparison with a Cardiologist

As part of the clinical study to validate the developed autonomous auscultation system, each participant was inspected by a resident cardiologist at Tygerberg Hospital. The cardiologist only had access to a stethoscope with which to formulate his diagnosis and did not view the ECG or echocardiograph of the participants prior to noting his findings. The findings of the cardiologist examinations are listed in Table A4 along with the participant details and echocardiograph diagnosis.

In comparison with the echocardiograph, the cardiologist diagnosed all the abnormal patients correctly as abnormal, but misdiagnosed one healthy participant. The cardiologist therefore scored a sensitivity of 100% and specificity of 96.6% in the clinical study. This is much better than the sensitivity and specificity obtained by the developed autonomous auscultation system (82% and 88% respectively). However, given the fact that the cardiologist has auscultation experience on likely several hundreds of patients and the autonomous system could only rely on the recordings of about 60 volunteers, significant improvement in the system's performance can be expected with the availability of more training data.

5.4 Chapter Summary

This chapter discussed the results obtained with the autonomous auscultation system in the clinical study at Tygerberg Hospital. Calculation of the sensitivity, specificity, predictive values and likelihood ratios as statistical performance measures is shown and a description of their significance is presented. The chapter concludes with a comparison between the performance of a cardiologist and the autonomous auscultation system.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 The Autonomous Auscultation System

The autonomous auscultation system developed in this study and used in concert with the PCG prototype device was able to successfully identify CVD in patients, with a sensitivity of 82% and specificity of 88%. Similarly, the results obtained with the Auscultation Jacket recordings, had sensitivity and specificity of ratings of 76.2% and 90.3%, respectively. The primary objective of this project was hereby achieved through the accurate and autonomous screening of patients for CVD in a clinical environment.

Through the use of Telemedicine, the success of this research project has far reaching implications in terms of providing 21st century healthcare to rural parts of Africa. It is envisioned that patient data could not only be collected by the PCG device for screening purposes, but also for remote consultation if a positive screening result is established. By transmitting patient data to specialists almost anywhere in the world, the system can aid in the early identification of life threatening cardiovascular diseases, which is vital to its effective treatment.

6.2 Data Acquisition

The heart sounds and ECG signals of male and female participants, which included 34 healthy persons and 28 patients with abnormal heart conditions, were successfully recorded at Tygerberg Hospital using the PCG device. The total number recorded participants amounted to 114 (including those previously recorded with the Auscultation Jacket) and were found to be adequate to train and validate the ensemble neural network classification system through implementation of the leave-one-out cross-validation method. The recording procedure, described in Appendix C, was accepted by all participants in the project and no negative reactions were encountered while performing recordings with the PCG device.

The PCG device was able to record good quality signals for most of the participants. However, for some obese participants the electronic stethoscopes did not make proper contact in the correct auscultation positions, and resulted in noisy recordings. Also, with these participants the device needed to be pressed against

the chest in order for the stethoscopes to sufficiently compress the chest fat under the skin. This was necessary for the stethoscopes to effectively measure the vibrations caused by the heart. Without this compression, the fat and the skin absorbed most of the vibrations making the heart sounds almost immeasurably faint. In some cases it would have been desirable to be able to extend the device along its length and have its centre part be flexible to better accommodate patients who suffered from obesity.

The ECG recording quality was excellent in many of the underweight abnormal and normal participants, but was generally poor for women and obese patients. This was mainly due to the fact that the limb leads of the ECG did not reach far enough towards the shoulders and abdomen of the patient, which creating noisy or faint traces (Lubbe, 2009). A possible solution would be to attach the four limb leads to extension cords, allowing them to be placed to the standard positions (ankles and wrists). The synchronisation of the ECG and heart sounds also posed a major challenge in this project. In future designs this can be solved by having an ECG sample through a computer's soundcard or a compact, multichannel USB soundcard built into the device. This will enable the recording software on the PC to synchronise all the signals automatically.

A future study could benefit if a comparison is not only made between the performance of the autonomous auscultation system and a cardiologist (Section 5.3), but with an otherwise qualified clinician such as a nurse or general practitioner as well. By comparing their auscultation performance to that of the autonomous auscultation system, the value of such a system could further be illustrated.

For the future success of a commercial automated auscultation system, an open EMR standard for storing and retrieving synchronised heart sound and ECG recordings will be essential. This standard will allow for much better collaboration between researchers around the world and, more importantly, aid in the availability of data, with which to test the performance of current and future systems.

6.3 Feature Extraction

Identification of significant features which would indicate the presence of CVD in heart sounds remains a subject that requires active research. It lies at the heart of the autonomous auscultation system and is critical to its screening ability.

The EEMD data analysis method proved useful in extracting significant features from heart sounds. Investigation of the method showed it has an exceptional ability to decompose heart sounds into meaningful components. Spectral analysis of these components through the application of auto regressive modelling yielded features which resulted in a significant screening improvement over the pathol-

ogy-based, time domain features. Unfortunately, due to time constraints posed on this study the spectral characteristics of the implemented, 12-order, auto regressive models could not be investigated further. But, due to the success of this method its future research is certainly warranted. Principal component analysis as well as the statistical overlap factor can also prove useful for providing information with regards to which auto regressive coefficients separate the various pathologies best.

Implementation of certain features from the ECG recordings has significant potential to increase the capabilities of the autonomous auscultation system. Examples of such ECG features that may prove useful are the height of the P-wave relative to the R-peak and the appearance of an inverted QRS-complex. A negative T-wave, the presence of arrhythmias and QRS-axis deviation could also be autonomously detected to enhance the screening capabilities of the system.

6.4 Classification

The ensemble ANN used in the final implementation of this study was able to achieve consistent cross-validation results through repeated training and simulation runs. The cross-validation results from Chapter 5 indicate that the classification system was able to adequately generalize over the input space.

The effects of using each cardiac cycle as an input to the classifier, and thereby having multiple multi-dimensional data points per participant, have been briefly investigated in this study. However, it was found that only in a few cases an improvement was obtained, but overall performance was compromised. Computing the median over all the cycles of each participant yielded better and more consistent results overall.

To further assist clinicians in the decision making process, a future version of the software could graphically portray the weights of the trained ANN. This could be beneficial as it would indicate how the network derives a particular result and not just appear to function as a black box.

The implementation of a single, unitary artificial neural network for classification was investigated, but the ensemble ANN repeatedly achieved superior results. The ensemble method also leaves much room for future investigation and further improvement. Increasing the number of ensemble members could potentially increase performance, although at the cost of computational speed. Different methods for combining the ensemble decisions also warrant investigation.

The addition of dynamic weights to the decisions of ensemble members could yield a significant improvement in future systems. Making the weights adjustable based on the certainty of each member's output, or whether a problematic output

is produced (Section 5.2), will form a critical part of future work along this line.

6.5 A note on Classification Performance and Hardware

Through the development of the current autonomous auscultation system, an issue of bias in the results was raised. By adjusting neural network parameters and choosing features based on classification performance, may have biased the system around the current data. Ideally, a separate dataset should have been kept which has not been used during the development of the system. This dataset is then used to test the system, after completion of its development, to determine its unbiased performance. However, due to the limited amount of available data, this was not practical in the current study. It is, however, strongly recommended for future development and testing of the system.

Combining the data recorded with the PCG device and the Auscultation Jacket into the same training and simulation datasets led to an extremely poor classification ability of the system. This phenomenon is attributed to the differences of the recording hardware. The fact that the positioning of the stethoscopes in the devices is not exactly similar, as well as the differences in the bell and diaphragm shapes and materials of the two devices, is thought to be the largest contributing factors.

The hypothesis that the autonomous auscultation system is sensitive to the physical properties of the recording hardware is further strengthened by the fact that when the datasets are evaluated separately, good classification performance was achieved, as reported in Chapter 5. This observation led to the conclusion that the system must be trained specifically for each type of recording hardware.

6.6 Achievement of the Project Objectives

In this section the achievement of the project objectives is summarised and the degree to which each of the objectives were obtained is discussed with regard to its original formulation.

6.6.1 Original project objectives

The objective of this research project was to develop software capable of detecting CVD from heart sounds recorded with electronic stethoscopes. In order to achieve this objective the following had to be accomplished:

4. Record heart sounds and electrocardiogram (ECG) data from patients with CVD as well as from healthy participants.
5. Research and implement suitable signal processing algorithms for the detection of anomalies in the recorded heart sounds which would indicate the presence of CVD.
6. Develop an automated classification scheme that can successfully partition the healthy participants from those with CVD.
7. Cross-validate the implemented classification scheme in a clinical study to determine its capability to screen future patients.
8. Investigate the strengths, weaknesses, feasibility and the acceptance of the recording hardware by patients and clinicians in a hospital environment.

6.6.2 Project achievements

The first objective involved a clinical study to assess the practical application and performance of the developed autonomous auscultation system. The clinical study was executed successfully and the recordings of 34 healthy participants and 28 patients with CVD could be used to validate the system against the gold standard echocardiograph. Each participant received the echo and was additionally examined by a trained cardiologist.

The purpose of the second objective was to investigate methods that would prove useful for identifying anomalies in the heart sounds of cardiac patients in an automated way. The EEMD data analysis method's application to heart sounds was investigated to this end. The research indicated that it provided a useful decomposition which can be used to identify signs of CVD in these sounds. The challenge of automating the analysis was solved by the use of correlation which could identify the significant components of the EEMD analysis. To characterise these components, the use of auto regressive modelling was investigated. The research showed that the combination of these signal processing algorithms successfully provided features which indicated the presence or absence of CVD in the heart sound recordings.

The third and fourth objectives were to develop a classification scheme that would identify which participants suffered from CVD and to cross-validate the classifier to accurately report its usefulness as a diagnostic screening tool. An ensemble neural network, discussed in Section 4.6, was developed to this end. The ensemble neural network was evaluated with the leave-one-out cross-validation method and proved to be a reliable screening tool by producing good cross-validation results in terms of the sensitivity, specificity and predictive values and

likelihood ratios.

The final objective of the project was to evaluate the PCG device in a clinical setting. This objective was achieved as part of the clinical study. Successful evaluation of the device was accomplished by enquiring and noting the opinions and experiences of the participants and clinicians involved in the study. The comfort and/or discomforts related to its application and the recording procedure was also investigated among the participants. Some of the physical recordings were examined by a cardiologist to obtain their impressions regarding the usefulness, quality and clinical relevance of the recordings. It was established that the device still requires development, specifically on improving the ECG recording quality and automatically synchronising it with heart sounds. Despite these minor shortcomings related to the device, all the participants and clinicians involved found it to be a practical tool and realised its potential value in basic medical facilities.

7. APPENDIX A: PARTICIPANT DATA

Table A1: PCG device test output (2009/10).

Participant	Echocardiograph		Ensemble decisions		Classification output			
ID #	Abn/Norm	Diagnosis	Class 1	Class 2	Class 1	Class 2	Abn/Norm	Confidence
3	N	Normal	87	12	1	0	N	87.9%
64	N	Normal	79	20	1	0	N	79.8%
39	N	Normal	66	33	1	0	N	66.7%
51	N	Normal	76	23	1	0	N	76.8%
8	N	Normal	81	18	1	0	N	81.8%
63	N	Normal	80	19	1	0	N	80.8%
37	N	Normal	79	20	1	0	N	79.8%
66	N	Normal	81	18	1	0	N	81.8%
67	N	Normal	84	15	1	0	N	84.8%
62	N	Normal	81	18	1	0	N	81.8%
15	N	Normal	74	25	1	0	N	74.7%
61	N	Normal	87	12	1	0	N	87.9%
68	N	Normal	81	18	1	0	N	81.8%
9	N	Normal	78	21	1	0	N	78.8%
10	N	Normal	82	17	1	0	N	82.8%
57	N	Normal	56	43	1	0	N	56.6%
12	N	Normal	80	19	1	0	N	80.8%
16	N	Normal	74	25	1	0	N	74.7%
69	N	Normal	74	25	1	0	N	74.7%
22	N	Normal	83	16	1	0	N	83.8%
24	N	Normal	82	17	1	0	N	82.8%
25	N	Normal	91	8	1	0	N	91.9%
26	N	Normal	86	13	1	0	N	86.9%
27	N	Normal	86	13	1	0	N	86.9%
30	N	Normal	67	32	1	0	N	67.7%
31	N	Normal	90	9	1	0	N	90.9%
32	N	Normal	87	12	1	0	N	87.9%
33	N	Normal	88	11	1	0	N	88.9%
34	N	Normal	91	8	1	0	N	91.9%
35	N	Normal	71	28	1	0	N	71.7%
38	N	Normal	16	83	0	1	A	83.8%
47	N	Normal	26	73	0	1	A	73.7%
65	N	Normal	9	90	0	1	A	90.9%
28	N	Normal	41	58	0	1	A	58.6%
14	A	AS,AR,MS,MR	39	60	0	1	A	60.6%
18	A	MR,AR	18	81	0	1	A	81.8%

Participant	Echocardiograph		Ensemble decisions		Classification output			
ID #	Abn/Norm	Diagnosis	Class 1	Class 2	Class 1	Class 2	Abn/Norm	Confidence
19	A	MR	19	80	0	1	A	80.8%
20	A	MS	19	80	0	1	A	80.8%
21	A	AR	21	78	0	1	A	78.8%
29	A	AS	34	65	0	1	A	65.7%
40	A	MS,MR,AR	15	84	0	1	A	84.8%
41	A	MR	40	59	0	1	A	59.6%
42	A	MR,AR,AS	9	90	0	1	A	90.9%
43	A	AR,AS	15	84	0	1	A	84.8%
44	A	AS	11	88	0	1	A	88.9%
45	A	VSD	16	83	0	1	A	83.8%
46	A	AS	40	59	0	1	A	59.6%
48	A	PS	19	80	0	1	A	80.8%
49	A	AS	19	80	0	1	A	80.8%
50	A	VSD	23	76	0	1	A	76.8%
52	A	MS	36	63	0	1	A	63.6%
53	A	AS	27	72	0	1	A	72.7%
55	A	MR	13	86	0	1	A	86.9%
56	A	AS	12	87	0	1	A	87.9%
58	A	MR	42	57	0	1	A	57.6%
59	A	MR,AR,AS	14	85	0	1	A	85.9%
4	A	MR	45	54	0	1	A	54.5%
23	A	AR	81	18	1	0	N	81.8%
36	A	MS	71	28	1	0	N	71.7%
54	A	prosthetic MV	61	38	1	0	N	61.6%
17	A	MR	91	8	1	0	N	91.9%
60	A	MS	57	42	1	0	N	57.6%

Table A2: Auscultation Jacket test output (2009/10).

	Echocardiograph		Ensemble decisions		Classification output			
Participant	Abn/Norm	Diagnosis	Class 1	Class 2	Class 1	Class 2	Abn/Norm	Confidence
1	N	Normal	85	14	1	0	N	85.9%
2	N	Normal	57	42	1	0	N	57.6%
3	N	Normal	96	3	1	0	N	97.0%
4	N	Normal	91	8	1	0	N	91.9%
5	N	Normal	94	5	1	0	N	94.9%
6	N	Normal	92	7	1	0	N	92.9%
7	N	Normal	98	1	1	0	N	99.0%
8	N	Normal	94	5	1	0	N	94.9%
9	N	Normal	89	10	1	0	N	89.9%
10	N	Normal	95	4	1	0	N	96.0%
11	N	Normal	94	5	1	0	N	94.9%
12	N	Normal	94	5	1	0	N	94.9%
13	N	Normal	87	12	1	0	N	87.9%
14	N	Normal	93	6	1	0	N	93.9%
15	N	Normal	95	4	1	0	N	96.0%
16	N	Normal	94	5	1	0	N	94.9%
17	N	Normal	93	6	1	0	N	93.9%
18	N	Normal	93	6	1	0	N	93.9%
19	N	Normal	97	2	1	0	N	98.0%
20	N	Normal	90	9	1	0	N	90.9%
21	N	Normal	92	7	1	0	N	92.9%
22	N	Normal	91	8	1	0	N	91.9%
23	N	Normal	96	3	1	0	N	97.0%
24	N	Normal	95	4	1	0	N	96.0%
25	N	Normal	98	1	1	0	N	99.0%
26	N	Normal	60	39	1	0	N	60.6%
27	N	Normal	89	10	1	0	N	89.9%
28	N	Normal	92	7	1	0	N	92.9%
29	N	Normal	19	80	0	1	A	80.8%
30	N	Normal	16	83	0	1	A	83.8%
31	N	Normal	9	90	0	1	A	90.9%
32	A	AS,MS,AR,MR	15	84	0	1	A	84.8%
33	A	MS,MR	13	86	0	1	A	86.9%
34	A	VSD,AR,PR	8	91	0	1	A	91.9%
35	A	AR	11	88	0	1	A	88.9%
36	A	AR	27	72	0	1	A	72.7%
37	A	AR	21	78	0	1	A	78.8%
38	A	AS,AR,MS	45	54	0	1	A	54.5%
39	A	MR,PR	23	76	0	1	A	76.8%
40	A	AR,MR	23	76	0	1	A	76.8%
41	A	MS,MR,AR	8	91	0	1	A	91.9%
42	A	AS,AR,MR	5	94	0	1	A	94.9%
43	A	MR	8	91	0	1	A	91.9%
44	A	MS,MR	17	82	0	1	A	82.8%
45	A	MR	21	78	0	1	A	78.8%
46	A	AS,MS,AR,MR	18	81	0	1	A	81.8%
47	A	AR	19	80	0	1	A	80.8%
48	A	MR	73	26	1	0	N	73.7%
49	A	VSD	89	10	1	0	N	89.9%
50	A	MR	64	35	1	0	N	64.6%
51	A	MR,AS,AR	91	8	1	0	N	91.9%
52	A	VSD	94	5	1	0	N	94.9%

Table A3: Auscultation Jacket participant #35. Example output of ensemble neural network decisions.

	Ensemble Neural Net Outputs		Neural Net Decisions	
Member	Class 1 (N)	Class 2 (A)	Class 1 (N)	Class 2 (A)
1	0.9677	0.9255	1	0
2	0.0166	0.9974	0	1
3	0.0017	0.9992	0	1
4	0.0130	0.9146	0	1
5	0.0043	0.9931	0	1
6	0.0000	1.0000	0	1
7	0.0005	0.9997	0	1
8	0.0013	0.9920	0	1
9	0.0066	0.9933	0	1
10	0.0000	1.0000	0	1
11	0.4938	0.9987	0	1
12	0.1101	0.9953	0	1
13	0.0032	0.9576	0	1
14	0.0090	0.6011	0	1
15	0.1539	0.9991	0	1
16	0.0008	0.9961	0	1
17	0.0000	1.0000	0	1
18	0.0033	0.9799	0	1
19	0.2143	0.9962	0	1
20	0.0183	0.9744	0	1
21	0.9813	0.9633	1	0
22	0.0036	0.9801	0	1
23	0.0000	1.0000	0	1
24	0.0092	0.9993	0	1
25	0.6178	0.6204	0	1
26	0.0000	1.0000	0	1
27	0.1618	0.9926	0	1
28	0.2251	0.9529	0	1
29	0.0022	0.9889	0	1
30	0.1849	0.9207	0	1
31	0.0000	1.0000	0	1
32	0.0000	0.9994	0	1
33	0.0029	0.9989	0	1
34	0.0107	0.0069	1	0
35	0.0187	0.9879	0	1
36	0.0042	0.9995	0	1
37	0.0000	1.0000	0	1
38	0.9977	0.0437	1	0
39	0.0002	0.9991	0	1
40	0.4813	0.8547	0	1
41	0.0001	0.8682	0	1
42	0.1375	0.9270	0	1
43	0.0000	1.0000	0	1
44	0.0011	0.9988	0	1
45	0.0016	0.8765	0	1
46	0.0052	0.9460	0	1
47	0.3578	0.8737	0	1
48	0.1357	0.5334	0	1
49	0.8260	0.9774	0	1
50	0.0340	0.9690	0	1
51	0.0458	0.9996	0	1
52	0.9986	0.9977	1	0
53	0.9953	0.0085	1	0
54	0.0000	1.0000	0	1
55	0.9793	0.0399	1	0

	Ensemble Neural Net Outputs		Neural Net Decisions	
Member	Class 1 (N)	Class 2 (A)	Class 1 (N)	Class 2 (A)
56	0.4506	0.9652	0	1
57	0.0017	0.9850	0	1
58	0.0009	0.9998	0	1
59	0.0015	0.6203	0	1
60	0.0006	0.9892	0	1
61	0.0007	0.9824	0	1
62	0.0609	0.9662	0	1
63	0.0006	1.0000	0	1
64	0.0001	0.0000	1	0
65	0.1074	0.1705	0	1
66	0.1594	0.9428	0	1
67	0.0102	0.5323	0	1
68	0.4734	0.7501	0	1
69	0.0000	0.9981	0	1
70	0.0000	1.0000	0	1
71	0.0492	0.9947	0	1
72	0.0262	0.9792	0	1
73	0.2570	0.8951	0	1
74	0.1570	0.9650	0	1
75	0.0423	0.9954	0	1
76	0.4506	0.6477	0	1
77	0.0001	0.9995	0	1
78	0.0058	0.9996	0	1
79	0.0101	0.0523	0	1
80	0.0033	0.9996	0	1
81	0.0940	0.9880	0	1
82	0.4449	0.6401	0	1
83	0.9864	0.8065	1	0
84	0.0000	1.0000	0	1
85	0.0451	0.9030	0	1
86	0.0367	0.6772	0	1
87	0.8494	0.0217	1	0
88	0.0000	1.0000	0	1
89	0.0000	1.0000	0	1
90	0.1303	0.4812	0	1
91	0.0000	0.0077	0	1
92	0.0000	0.9999	0	1
93	0.0197	0.8622	0	1
94	0.1063	0.9985	0	1
95	0.1445	0.0017	1	0
96	0.0012	0.9985	0	1
97	0.0560	0.9709	0	1
98	0.0002	1.0000	0	1
99	0.0361	0.9910	0	1
		Total Decisions:	**11**	**88**

Table A4: Participants recorded with the PCG device at Tygerberg Hospital (2008/06 - 2009/06).

								Cardiologist examination						Echocardiograph
ID #	Age	Gender 0=♂ 1=♀	Weight Kg	Height cm	BMI Kg/m^2	Chest cm	Waist cm	Apex Beat	S1	S2	Extra sounds	Systolic	Diastolic	N=0 ABN=1	
1	39	0	89	172	30.1	103	96	6 ICS MCL	normal	normal	0	0	0	0	normal
2	30	0	93	172	31.4	112	97	6 ICS MCL	normal	normal	0	0	0	0	normal
3	32	0	79	180	24.4	93	86	6 ICS MCL	normal	phys. Split	0	0	0	0	normal
4	46	0	62	168	22.0	89	82	6 ICS lat MCL	loud	normal	0	PSM & ESM	0	1	MR
5	32	0	77	175	25.1	100	81	6 ICS MCL	normal	normal	0	0	0	0	normal
6	31	0	58	178	18.3	83	73	5 ICS MCL	normal	phys. Split	0	PSM	0	1	MR
7	33	0	107	193	28.7	110	84	5 ICS MCL	normal	normal	0	0	0	0	normal
8	21	0	62	168	22.0	88	82	5 ICS MCL	normal	normal	0	0	0	0	normal
9	24	0	58	164	21.6	85	78	5 ICS MCL	normal	normal	0	0	0	0	normal
10	22	0	73	177	23.3	86	87	5 ICS MCL	normal	normal	0	0	0	0	normal
11	27	0	51	175	16.7	85	72	6 ICS MCL	normal	loud	0	PSM	0	1	MR
12	26	0	79	175	25.8	101	87	5 ICS MCL	split	normal	0	0	0	0	normal
13	24	0	87	183	26.0	96	88	5 ICS MCL	normal	normal	0	0	0	0	normal
14	59	1	52	168	18.4	76	72	7 ICS AAL	soft	normal	0	ESM & PSM	Decrescendo	1	AS,AR,MS,MR
15	23	0	65	180	20.1	89	77	5 ICS MCL	normal	normal	0	0	0	0	normal
16	23	0	70	172	23.7	92	79	5 ICS MCL	split	normal	0	0	0	0	normal
17	28	1	46	169	16.1	72	60	5 ICS MCL	loud	loud	0	PSM	0	1	MR
18	57	1	76	160	29.7	90	92	7 ICS MCL	normal	normal	0	PSM	0	1	MR,AR
19	76	1	56	156	23.0	86	92	5 ICS MCL	normal	normal	0	ESM	0	1	MR
20	61	0	80	173	26.7	104	102	5 ICS MCL	loud	normal	0	0	Decrescendo	1	MS
21	46	0	64	162	24.4	100	75	5 ICS MCL	loud	normal	0	0	Decrescendo	1	AR
22	29	0	100	194	26.6	110	98	not felt	normal	normal	0	0	0	0	normal
23	36	0	45	165	16.5	86	68	7 ICS lat MCL	normal	normal	0	0	Decrescendo	1	AR
24	32	0	87	181	26.6	105	91	5 ICS MCL	normal	normal	0	0	0	0	normal
25	36	1	55	163	20.7	85	81	6 ICS MCL	normal	normal	0	0	0	0	normal
26	48	1	50	161	19.3	80	76	6 ICS MCL	normal	normal	soft rub	PSM	0	1	normal
27	29	1	74	165	27.2	100	92	not felt	normal	normal	0	0	0	0	normal
28	33	0	81	178	25.6	106	86	5 ICS MCL	normal	normal	0	0	0	0	normal
29	79	0	67	169	23.5	93	86	not felt	soft	normal	0	ESM	0	1	AS
30	42	1	73	166	26.5	92	94	5 ICS MCL	normal	phys. Split	0	soft PSM	0	0	normal
31	53	1	89	154	37.5	115	109	5 ICS MCL	normal	normal	0	0	0	0	normal
32	43	1	68	175	22.2	99	87	5 ICS MCL	split	normal	0	0	0	0	normal
33	39	1	73	171	25.0	92	97	5 ICS MCL	normal	normal	0	0	0	0	normal
34	50	1	68	164	25.3	100	96	5 ICS MCL	normal	phys. Split	0	0	0	0	normal
35	47	1	80	152	34.6	109	115	not felt	normal	normal	0	0	0	0	normal
36	29	0	54	167	19.4	81	72	6 ICS AAL	loud	loud	OS	0	rumble	1	MS
37	51	1	75	154	31.6	101	97	5 ICS MCL	normal	normal	0	0	0	0	normal
38	41	1	72	179	22.5	91	91	5 ICS MCL	normal	normal	0	soft ESM	0	0	normal
39	49	1	99	160	38.7	122	120	5 ICS MCL	normal	normal	0	0	0	0	normal
40	45	1	109	161	42.1	115	128	not felt	loud	normal	0	PSM & ESM	0	1	MS,MR,AR
41	15	0	61	164	22.7	75	65	5 ICS lat MCL	loud	normal	0	PSM & ESM	0	1	MR
42	33	0	56	181	17.1	87	82	5 ICS lat MCL	normal	loud	0	PSM & ESM	Decrescendo	1	MR,AR,AS
43	52	0	52	167	18.6	86	79	7 ICS lat MCL	normal	normal	rub	PSM & ESM	Decrescendo	1	AR,AS
44	31	0	85	175	27.8	95	83	5 ICS MCL	normal	normal	0	Early SM	0	1	AS
45	23	1	51	157	20.7	88	82	5 ICS MCL	loud	normal	0	PSM	0	1	VSD

ID #	Age	Gender 0=♂ 1=♀	Weight Kg	Height cm	BMI Kg/m^2	Chest cm	Waist cm	Cardiologist examination						Echocardiograph
								Apex Beat	S1	S2	Extra sounds	Systolic	Diastolic	N=0 ABN=1	
46	42	1	69	159	27.3	91	97	5 ICS MCL	normal	normal	ejection click	ESM	0	1	AS
47	22	1	47	152	20.3	79	78	5 ICS MCL	normal	normal	0	0	0	0	normal
48	22	1	51	155	21.2	83	80	5 ICS MCL	soft	loud	0	ESM	0	1	PS
49	77	0	90	171	30.8	111	109	not felt	normal	soft	0	ESM	0	1	AS
50	43	0	98	176	31.6	100	96	5 ICS MCL	normal	normal	0	PSM	0	1	VSD
51	53	0	82	169	28.7	103	97	5 ICS MCL	loud	normal	0	very soft PSM	0	0	normal
52	25	1	62	162	23.6	85	79	7 ICS MCL	loud	normal	OS	0	rumble	1	MS
53	74	0	84	177	26.8	106	99	not felt	normal	normal	0	ESM	0	1	AS
54	41	0	73	180	22.5	95	92	not felt	soft	prosthetic click	0	midsystolic	0	1	prosthetic mv
55	32	1	84	164	31.2	115	109	5 ICS MCL	normal	normal	0	PSM	0	1	MR
56	63	0	119	165	43.7	127	122	5 ICS MCL	normal	normal	0	ESM	0	1	AS
57	35	0	85	182	25.7	100	96	5 ICS MCL	normal	normal	0	0	0	0	normal
58	43	0	62	167	22.2	88	85	5 ICS MCL	normal	normal	0	PSM	0	1	MR
59	58	1	128	162	48.8	128	133	not felt	loud	normal	0	ESM	Decrescendo	1	MR,AR,AS
60	13	1	63	155	26.2	92	93	not felt	normal	loud	OS	0	rumble	1	MS

8. APPENDIX B: CONSENT FORM

PARTICIPANT INFORMATION LEAFLET AND CONSENT FORM

TITLE OF THE RESEARCH PROJECT:
Recognition of Cardiac Abnormalities with the Aid of the Precordial Auscultation Device

REFERENCE NUMBER:

PRINCIPAL INVESTIGATOR: Dr WW Lubbe

ADDRESS: Department of Internal Medicine

University of Stellenbosch and Tygerberg Academic Hospital Complex

PO Box 19063, Tygerberg, 7505, South Africa

CONTACT NUMBER: 021 9384400

You are being invited to take part in a research project. Please take some time to read the information presented here, which will explain the details of this project. Please ask the study staff or doctor any questions about any part of this project that you do not fully understand. It is very important that you are fully satisfied that you clearly understand what this research entails and how you could be involved. Also, your participation is **entirely voluntary** and you are free to decline to participate. If you say no, this will not affect you negatively in any way whatsoever. You are also free to withdraw from the study at any point, even if you do agree to take part.

This study has been approved by the **Committee for Human Research at Stellenbosch University** and will be conducted according to the ethical guidelines and principles of the international Declaration of Helsinki, South African Guidelines for Good Clinical Practice and the Medical Research Council (MRC) Ethical Guidelines for Research.

What is this research study all about?

Description of study

A precordial (meaning chest) auscultation device has been developed by an engineer in the Mechanical Engineering department of Stellenbosch University. The device is being designed to tell the difference between normal and abnormal heart sounds and it can also make an ECG (measures the electrical activity of the heart).

One of the intended purposes of the precordial cardiogram is to use it in the communities to identify abnormal hearts and see which patients need to be referred to a specialist, because many areas don't have specialized medical staff working there.

Details of the procedure

You will be examined by a cardiologist and an echocardiogram will be done. You will also have a standard 12-lead ECG done. For this you will have to remove your shirt. The precordial auscultation device will then be placed on your chest. Before placing the device on you, standard medical electrode gel will be applied to each stethoscope as this results in better recordings. The electrode gel might feel a bit cold at first but this will not last long.

All procedures are non-invasive and completely painless. You will be asked to lie on your back, the bed will be tilted at 30 degrees, breathe normally and a recording will be done for 10 seconds. You will then be asked to remain on your back, hold your breath in expiration for 5 seconds while the recording is being done. A recording will be done with you sitting upright and holding your breath in expiration for 5 seconds. You will also be asked to sit upright, bend slightly forward and a 10 second recording will be done. The final recording will be done while you lie on your left side and hold your breath in expiration for 5 seconds. All procedures will be performed on the 8th floor in Tygerberg Hospital at the cardiology clinic.

The precordial auscultation device itself (technical details)

The precordial auscultation device is fitted with 4 electronic stethoscopes and a 12-lead ECG and consists of 1 plastic front piece. The stethoscopes are used to "listen" to your heart. To obtain recordings the device is simply held in place over your chest. The 4 stethoscopes are embedded in the plastic and when placed over your chest will cover the same areas that a doctor will listen to during an examination. The stethoscopes are made from metal.

The ECG electrodes are also embedded in the plastic and they obtain a 12 lead ECG. An ECG measures the electrical activity in the heart and many disease processes of the heart are diagnosed with ECG's, like a heart attack.

All the stethoscopes as well as the ECG are USB-powered. They plug into a computer where the information is recorded and the computer will comment on the heart sounds.

Why have you been invited to participate?

You have been invited to participate because you have a heart murmur that can be heard with a stethoscope. We would like to see if the precordial auscultation device can also detect the murmur.

Will you benefit from taking part in this research?

You will not benefit directly from this study, but by participating may help us in detecting people with similar heart murmurs so that they may receive treatment.

Are there any risks involved in your taking part in this research?

There are no risks involved in taking part in this research.

Who will have access to your medical records?

The information collected will be treated as confidential and protected. If it is used in a publication or thesis, your identity will remain anonymous.

Will you be paid to take part in this study and are there any costs involved?

There will be no costs involved for you, if you do take part.

Declaration by participant

By signing below, I ..... agree to take part in a research study entitled: Recognition of Cardiac Abnormalities with the Aid of the Precordial Auscultation Device.

I declare that:

- I have read or had read to me this information and consent form and it is written in a language with which I am fluent and comfortable.
- I have had a chance to ask questions and all my questions have been adequately answered.
- I understand that taking part in this study is **voluntary** and I have not been pressurised to take part.

Signed at (*place*) ..... on (*date*) ..... 2005.

....

Signature of participant **Signature of witness**

Declaration by investigator

I *(name)* ..... declare that:

- I explained the information in this document to .....
- I encouraged him/her to ask questions and took adequate time to answer them.

- I am satisfied that he/she adequately understands all aspects of the research, as discussed above
- I did/did not use a translator. (*If a translator is used then the translator must sign the declaration below.*

Signed at (*place*) ..... on (*date*) ..... 2005.

....
Signature of investigator **Signature of witness**

Declaration by translator

I *(name)* ..... declare that:

- I assisted the investigator (*name*) ..... to explain the information in this document to (*name of participant*) ..... using the language medium of Afrikaans/Xhosa.
- We encouraged him/her to ask questions and took adequate time to answer them.
- I conveyed a factually correct version of what was related to me.
- I am satisfied that the participant fully understands the content of this informed consent document and has had all his/her question satisfactorily answered.

Signed at (*place*) ..... on (*date*) ..... 2005.

....
Signature of translator **Signature of witness**

9. APPENDIX C: RECORDING PROCEDURE

The procedure that was used to record participants' data at Tygerberg Hospital for this study was as follows:

1. The volunteer is informed about the study, its purpose and the procedure for performing recordings.
2. The participant signs the consent form.
3. The participant is instructed to lie on his/her back with the headrest of the bed inclined at 30°.
4. Electrolyte gel is placed on the ECG electrodes
5. The PCG device is placed on the participant's chest with the stethoscopes at the auscultation positions.
6. The ECG and heart sound recordings are initiated on the PC.
7. The device is tapped to create intentional artefacts in the signals.
8. Record for 30 seconds.
9. Instruct the participant to take a deep breath, exhale and not to inhale for as long as possible.
10. Record for 30 seconds after the participant has again inhaled.
11. Save the audio recording as a pulse-code modulated wave file and the ECG in a text file on the personal computer.
12. Repeat 3 more times from step 4.
13. A separate 12-lead ECG is recorded by the clinician.
14. The participant is examined by the cardiologist with only a stethoscope.
15. The participant is examined by an echocardiograph.

10. APPENDIX D: CARDIOLOGIST FORM

EXAMINATION BY CARDIOLOGIST

Participant number: _____

Gender: _____

Age: _____

Height: _____

Weight: _____

Chest Circumference: _____

Blood pressure: _____

Apex beat position: _____

1st heart sound: _____

2nd heart sound: _____

Additional sounds: _____

Systolic murmurs: _____

Diastolic murmurs: _____

11. APPENDIX E: SOUNDCARD ECG

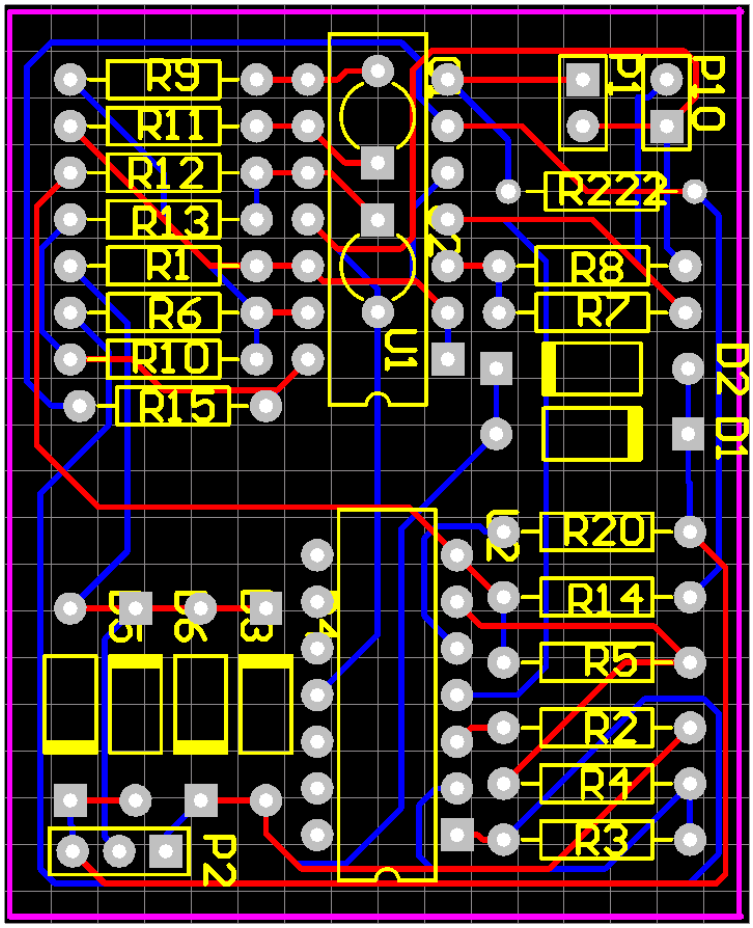

Figure A1: Soundcard ECG printed circuit board layout with component overlay.

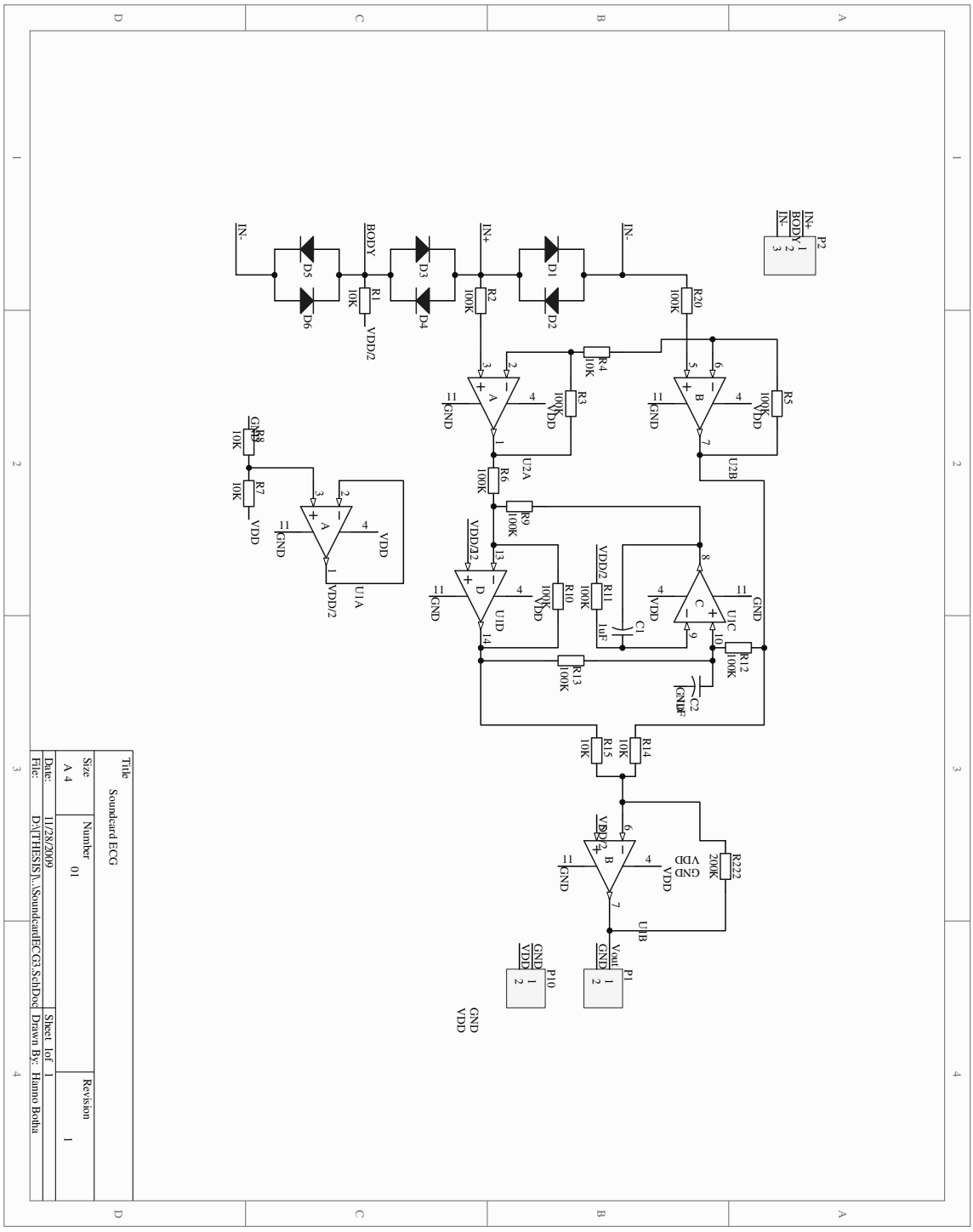

Figure A2: Soundcard ECG circuit diagram.

12. REFERENCES

Altman, D. & Bland, J.M., 1994. Sttistics Notes: Diagnostic tests 2: Predictive values. *British Medical Journal*, 309, p.102.

Andrisevic, N., Ejaz, K., Gutierrez, F.R. & Flores, R.A., 2005. Detection of Heart Murmurs Using Wavelet Analysis and Artificial Neural Networks. 127, pp.899 - 905.

Ari, S. & Saha, G., 2008. Classification of heart sounds using empirical mode decomposition based features. *International Journal of Medical Engineering and Informatics*, 1, pp.91 - 107.

Bhatikar, S.R., DeGroff, C. & Mahajan, R.L., 2004. A classifier based on the artificial neural network approach for cardiologic auscultation in pediatrics. *Artificial Intelligence in Medicine*, 33, pp.251 - 260.

Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. New York: Springer.

Burke, M.J. & Nasor, M., 2002. The time relationships of the constituent components of the human electrocardiogram. *Journal of Medical Engineering & Technology*, 26, pp.1 - 6.

Cathers, I., 1995. Neural Network assisted Cardiac Auscultation. *Artificial Intelligence in Medicine*, 7, pp.53-66.

Charleston, V.S., Aljama, C.A. & Gonzales, C.R., 2006. Analysis of Simulated Heart Sounds by Intrinsic Mode Ffunctions. In *Proc. 28th IEEE EMBS Annual Int. Conf.* New York, USA, 2006.

Chauhan, S., Wang, P., Lim, C.S. & Anantharaman, V., 2008. Acomputer-aided MFCC-based HMMsystem for automatic auscultation. *Computers in Biology and Medicine*, 38, pp.221 - 223.

de Vos, J., 2004. *Automated Pediatric Cardiac Auscultation*. Masters Thesis. Stellenbosch, South Africa: University of Stellenbosch.

Debbal, S.M. & Bereksi, R.F., 2004. Heart beat sound analysis with the wavelet transform. *Journal of Mechanics in Medicine and Biology*, 4, pp.133 - 141.

Deeks, J.J. & Altman, D.G., 2004. Diagnostic tests 4: likelihood ratios. *British Medical Journal*, 329, pp.168 - 169.

Dietterich, T.G., 2000. Ensemble Methods for Machine Learning. In J. Kittler & F. Roli, eds. *Multiple Classifier Systems*. Berlin Heidelberg: Springer. pp.1-15.

Echeverria, J.C., Crowe, J.A., Woolfson, M.S. & Hayes-Gill, B.R., 2001. Application of empirical mode decomposition to heart rate variability analysis. *Journal of Medical & Biological Engineering & Computing*, 34, pp.471 - 479.

Flandrin, P., Rilling, G. & Gonçalvès, P., 2004. Empirical mode decomposition as a Filter Bank. *IEEE Signal Processing Letters*, 11, pp.112 - 114.

Gamero, L.G. & Watrous, R., 2003. Detection of the First and Second Heart Sound Using Probabilistic Models. In *Proceedings of the 25' Annual Intemational Conference of the IEEE EMBS.*, 2003.

Gilbert, D. & Tan, A.C., 2003. Ensemble machine learning on gene expression data for cancer classification. *Applied Bioinformatics*, 3, pp.75 - 83.

Greenhalgh, T., 1997. How to read a paper: Papers that report diagnostic or screening tests. *British Medical Journal*, 315, pp.540 - 543.

Guo, Z. et al., 1994. Artificial neural networks in computer-assisted classification of heart sounds in patients with porcine bioprosthetic valves. *Medical and Biological Engineering and Computing*, 23, pp.311 - 316.

Gupta, C.N., Palaniappan, R., Swaminathan, S. & Krishnan, S.M., 2005. Neural network classification of homomorphic segmented heart sounds. *Applied Soft Computing*, pp.286 - 297.

Huang, N.E. et al., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Procedures of the Royal Society of London*, 454, pp.903 - 995.

Huiying, L., Sakari, L. & Iiro, H., 1997. A heart sound segmentation algorithm using wavelet decomposition and reconstruction. *Computers in Cardiology*, pp.1630 - 1663.

Human, n.d. *Human*. [Online] Available at: http://www.humans.be/image/sites [Accessed 8 July 2008].

James, C.J. et al., 1997. Multireference adaptive noise canceling applied to the EEG. *IEEE Transactions on Biomedical Engineering*, 44, pp.775 - 779.

Klabunde, R., 2009. *Cardiovascular physiology concepts*. [Online] Available at: http://www.cvphysiology.com [Accessed 26 August 2009].

Koekemoer, D., 2008. Correspondence, 1 September. Stellenbosch, Western Cape, South Africa.

Koekemoer, H.L. & Scheffer, C., 2008. Heart Sound and Electrocardiogram Recording Devices for Telemedicine Environments. In *Proceedings of the 30th Annual International Conference of the IEEE EMBS*., 2008.

Leung, T.S. et al., 2000. Classification of heart sounds using time-frequency method and artificial neural networks. In *Proceedings of the 22"d Annual EMBS International Conference*., 2000.

Liang, H., Lukkarinen, S. & Hartimo, I., 1997. Heart Sound Segmentation Algorithm Based on Heart Sound Envelolgram. *Computers in cardiology*, 24, pp.105 - 108.

Lindberg, A.B., Weinrauch, L.A. & Zieve, D., 2009. *US National Library of Medicine*. [Online] Available at: http://www.nlm.nih.gov/medlineplus/ [Accessed 01 March 2009].

Lindberg, A.B., Weinrauch, L.A. & Zieve, D., n.d. *Medline Plus*. [Online] Available at: http://www.nlm.nih.gov/medlineplus/ [Accessed 01 March 2009].

Liu, B., Ciu, Q., Jiang, T. & Ma, S., 2004. A combinational feature selection and ensemble neural network method for classification of gene expression data. *BMC Bioinformatics*, 5, p.136.

Lubbe, W., 2009. *MD*. Correspondence, 3 March. Tygerberg Hosptal, Western Cape, South Africa.

Lubbe, W., 2009. *Recognition of Cardiac Abnormalities with the Aid of a Precordial Auscultation Device*. Mmed dissertation. Stellenbosch University.

Malarvili, M.B., Kamarulafizam, I., Hussain, S. & Helmi, D., 2003. Heart Sound Segmentation Algorithm Based on Instantaneous Energy of Electocardiogram. *Computers in Cardiology*, 30, pp.327 - 330.

Martini, F.H. & Bartholomew, E.F., 2003. *Essentials of Anatomy & Physiology*. New Jersey: Prentice Hall.

Messer, S.R., Agzarian, F. & Abbott, D., 2001. Optimal wavelet denoising for phonocardiograms. *Microelectronics Journal*, 4, pp.931 - 941.

Ministry of Health, 2009. *South African Department of Health Annual Reports*. [Online] Available at: http://www.doh.gov.za/docs/reports/index.html [Accessed 01 December 2009].

Mittra, A.K., Shukla, A. & Zadgaonkar, A.S., 2007. System simulation and comparative analysis of foetal heart sound de-noising techniques for advanced phonocardiography. *International Journal of Biomedical Engineering and Technology*, 1, pp.73-85.

Nguyen, D. & Widrow, B., 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. *Proceedings of the IJCNN*, 3, pp.21 - 26.

Nicogossian, A.E., Pober, D.F. & Roy, S.A., 2001. Evolution of Telemedicine in the Space Program and Earth Applications. *Telemedicine Journal and e-Health*, 7, pp.1-15.

Ning, J., Atanasov, N. & Ning, T., 2009. Quantitative Analysis of Heart Sounds and Systolic Heart Murmurs Using Wavelet Transform and AR Modeling. In *Proc. 31st Annual Int. Conf. of the IEEE EMBS*. Minneapolis, USA, 2009.

OSU Medical Centre, 2008. *Arrythmias*. [Online] Available at: http://medicalcenter.osu.edu/images/gerystone/ei_0018.jpg [Accessed 16 November 2008].

Pestana, J., Steyn, K., Leiman, A. & Hartzenberg, G., 1996. The direct and indirect costs of cardiovascular disease in South Africa. *South African Medical Journal*, 86, pp.679 - 684.

Proakis, J.G. & Manolakis, D.G., 2007. *Digital Signal Processing*. 4th ed. New Jersey: Prentice Hall.

Rangayyan, R.M., 2002. *Biomedical signal analysis: A case study approach*. New York: Wiley-Interscience.

Reed, T.R., Reed, N.E. & Fritzson, P., 2004. Heart Sound Analysis for Symptom Detection and Computer-aided Diagnosis. *Simulation Modelling Practice and Theory*, 12, pp.129 - 146.

Sengur, A. & Turkoglu, I., 2008. A hybrid method based on artificial immune system and fuzzy k-NN algorithm for diagnosis of heart valve diseases. *Expert Systems with Applications*, 35, pp.1011 - 1020.

Sharkey, A.J., 1996. On Combining Artificial Neural Nets. *Connection Science*, 8, pp.299 - 314.

Sinha, R.K., Aggarwal, Y. & Das, B.N., 2007. Backpropagation Artificial Neural Network Classifier to Detect Changes in Heart Sound due to Mitral Valve Regurgitation. *Medical Systems*, 31, pp.205 - 209.

Steyn, K., 2007. *Heart disease in South Africa*. [Online] Available at: http://www.heartfoundation.co.za/docs/heartmonth/HeartDiseaseinSA.pdf [Accessed 30 June 2008].

Steyn, A.G., Smit, C.F., Du Toit, S.H. & Strasheim, C., 2004. *Moderne Statistiek vir die Praktyk*. 6th ed. Pretoria: J.L.van Schaik.

Storrow, A., Lindsel, C., Peacock, W. & Collins, S., 2004. Computerized Detection of Third Heart Sounds Improves Sensitivity for the Emergency Department Diagnosis of Heart Failure. *Annals of emergency medicine*, Oktober. p.4.

Torry, J.N. & Mood, H.A., 1995. A Sub-Band Energy Tracking Algorithm for Heart Sound Segmentation. *Computers in Cardioloy*, pp.501 - 504.

Turkoglu, I., Arslan, A. & Ilkay, E., 2002. An expert system for diagnosing heart valve diseases. *Expert systems with Applications*, 23, pp.229 - 236.

Vanderplaats, G.N., 2005. *Numerical optimization techniques for engineering design*. 4th ed. Colorado Springs: Vanderplaats research and development.

Visagie, C., 2007. *Screening for abnormal heart sounds and murmurs by implementing neural networks*. Master's thesis. University of Stellenbosch, Stellenbosch, South Africa.

Wu, Z. & Haung, N.E., 2009. Ensemble Empirical mode decomposition: A Noise Assisted Data Analysis Method. *Advances in Adaptive Data Analysis*, 1, pp.1 - 41.

Zhou, Z., Jiang, Y. & Chen, S., 2002. Lung cancer cell identification based on artificial neural network ensembles. *Artificial inteligence in medicine*, 24, pp.25 - 36.